

VISIT...

LANZAROTE
Caliente.COM

Pharmacogenetic Testing

Wendell W. Weber

University of Michigan, Ann Arbor, USA

Maureen T. Cronin

ACLARA Biosciences, Mountain View, USA

1 Introduction	2
1.1 Historical Highlights	2
1.2 The Emergence of Pharmacogenetics	3
1.3 The Scope of Pharmacogenetics	3
1.4 The Pharmacological Profile of Human Drug Response	4
1.5 The Genetic Profile of Human Drug Response	5
1.6 The Rationale and Aims of Pharmacogenetics	5
1.7 Testing Pharmacogenetic Hypotheses	6
1.8 Summary	6
2 The Changing Scene	7
2.1 Human Protein Diversity	7
2.2 Some Biochemical Generalizations	7
2.3 DNA Polymorphism: Some Practical Matters	7
2.4 Molecular Heterogeneity of Pharmacogenetic Interest	8
2.5 The Effect of Mutations on the Gene Product	10
2.6 Summary	11
3 Genetic Profiling Using DNA Microarrays	11
3.1 The Origin of DNA Microarrays	11
3.2 DNA Microarray Capabilities and Limitations	13
3.3 Predictive Testing with DNA Microarray Assays	15
3.4 Emerging Applications of Microarray Technologies	16
3.5 The Future of Microarrays and Allied Technology	17
3.6 Summary	18
4 The Outlook for Genomics in Predictive Medicine	18
4.1 Options for Treating Genetic Disorders	18
4.2 Many Pharmacogenetic Polymorphisms are Exceptions to Experience with Single Gene Disorders	18

4.3 Emerging Applications of Pharmacogenetic Interest	19
4.4 Linking Human Genotyping to Clinical Applications	20
4.5 Summary	20
5 Summary: Progress in Pharmacogenetic Testing	20
Abbreviations and Acronyms	21
Related Articles	21
References	21

Drugs are administered to patients with the intention of achieving a planned therapeutic response. Yet drug prescription is a medical art because of the wide variation in individual responses to standard drug doses. In the 1960s, the field of pharmacogenetics emerged from the dedicated work of a small group of investigators who began to enrich the discipline of pharmacology by integrating it with human genetics. Pharmacogenetics provides the experimental framework to understand variability in human reactions to drugs and other exogenous substances as a function of intrinsic human genetic variability. The maturation of the human genome initiative has provided a wealth of primary genetic information and functional genomic data to fuel the understanding of genetic polymorphism and its functional consequences. Today, the field of pharmacogenetics is a well-integrated, worldwide network engaging a vast community of academic and industrial scientists.

One hallmark of pharmacogenetics studies is the large amount of genetic data that must be accumulated and integrated for high-resolution drug-response genotyping and subsequent phenotype profiling. This need for highly parallel genetic analysis has, in turn, fueled a demand for technical innovation to provide the tools necessary for its execution. The response has been a flurry of inventions in microfluidics and nanotechnology, some of which are scaled down versions of existing technologies, such as capillary electrophoresis and mass spectrometry, and some are completely new inventions. Outstanding in this technical revolution are DNA microarrays. Array technology has emerged as the most versatile and widely applied tool to support pharmacogenetics, for several reasons: arrays have the largest and most scalable capacity for parallel analysis, their inherent flexibility permits them to be used as independent analysis tools or as integrated components of more complex microfluidic systems and, finally, a wide variety of methods can be used to design and fabricate them. Pharmacogenetics is experiencing a period of rapid growth and definition. Arrays have proven themselves to be a technology capable of responding to the growing and changing needs of the current research

environment and are likely to be equally important as pharmacogenetics moves into the clinical arena.

1 INTRODUCTION

Drugs exert characteristic, reproducible effects when administered to individuals, and physicians have come to rely on them for the specific treatment of many human disorders. Most readers need not be convinced of the societal importance of these agents, but patients and physicians alike know that one individual may respond differently from another given the same dose of a drug. Such person-to-person variation in drug response is not a trivial matter. The analysis of the causes of natural variation in human drug response, the central problem of pharmacogenetics, requires a basic understanding of human pharmacology and human genetics.

The initial researches of a small cadre of dedicated investigators saw the emergence of pharmacogenetics in the 1960s as a new field of experimental science primarily concerned with reactions to drugs and other exogenous substances that occur because of a person's distinctive genetic makeup. The field has grown, particularly within the period from the late 1980s, into a well-integrated, worldwide network engaging a vast community of academic and industrial scientists with expertise in a broad spectrum of basic and clinical biomedical disciplines. Until recently, pharmacogenetic studies were usually limited to a few individuals and families in a research setting, but now such an investigation may be directed toward a large fraction of the human genome and involve simultaneous testing of multiple loci in many patients and healthy subjects. Such studies may, at times, be referred to as pharmacogenomics.

The establishment of laboratories dedicated to such large-scale pharmacogenetic testing is rapidly gaining momentum. This action has the attention of physicians who are in the front line of patient care, of those who are responsible for clinical laboratories and of academic and pharmaceutical scientists who are engaged in the discovery and clinical trial of new drugs. Because the genetics and the molecular basis of many pharmacogenetic traits are well established, because advance testing holds out high promise of improved patient care, and because the high potential for acceptance by the biomedical community, pharmacogenetics is in a strong position to link human genotyping and phenotyping rapidly with clinical application.

1.1 Historical Highlights

In an address before the British Medical Association in 1914,⁽¹⁾ Archibald Garrod proposed that enzymes were

somehow implicated in the detoxification of exogenous substances. During the intense flurry of research stimulated by the rediscovery of Mendel's laws of heredity around 1900, Cuenot, from studies of mice, and Garrod, from human studies, anticipated the connection of enzymes ("diastases") with the genetic material. Observations of physiological chemists, intrigued by the fate of chemicals in human subjects, had shown by this time that most drugs were excreted in forms that differed from those that were ingested. These observations, and Garrod's observations pertaining to a case of porphyria brought on by the hypnotic drug sulfonal, led Garrod to conclude that the ability of individuals to transform drugs into nontoxic conjugates, such as hippurates and glucuronates, protected them from the poisonous effects of these agents. Garrod was thus far ahead of his contemporaries in attributing unexpected drug responses of individuals to failure of their enzymes to detoxify these substances.

These ideas regarding the detoxification of chemicals surfaced again and again in Garrod's writings and teachings until the end of his life in the 1930s. He observed that substances in foods, in certain drugs, and in exhalations of animals and plants might produce effects in some persons wholly out of proportion to any they produce in most persons.⁽²⁾ During the 1920s and 1930s, others began to scrutinize person-to-person differences in the perception of odor and taste.⁽³⁾ For example, studies of "taste blindness" to the bitter-tasting substance *p*-ethoxyphenylurea was demonstrated to be a Mendelian character transmitted from parents to children as an autosomal trait. These deficits in sensory perception were the first indication of the high order of specificity to be expected in human response to chemicals, and the first to establish the heritable nature of such responses.⁽⁴⁾

By the time Garrod presented his ideas before the British Medical Association, chemists had identified virtually every major type of conjugation reaction that we know today. The exact dates of those discoveries is less relevant than the fact that they preceded by about 50 years the discovery of the other major category of detoxifying enzymes (P450s), commonly known as the microsomal enzymes or the drug-metabolizing enzymes. R. Tecwyn Williams, a pioneering pharmacologist at St. Mary's Medical School in London, saw evidence for a new, far-reaching principle of drug metabolism in the studies of the microsomal and conjugating enzymes, namely that the metabolic disposition of drugs in humans and animals occurs in two phases: first, oxidation, reduction, and hydrolysis, then conjugation, designating them as phase I and phase II metabolism respectively.

1.2 The Emergence of Pharmacogenetics

Until around 1950, progress slowed, but then several breakthroughs foreshadowing important events to come were made. In 1948, hemoglobinopathies S and C were identified and their mode of inheritance was determined. Hemoglobin S, the first hereditary protein variant to be identified, was also the first example of a variant protein in which replacement of a single amino acid residue within the protein molecule could be identified unequivocally with a change in the functional effect of the protein. In 1953, the double helix of DNA was described. A few years later, the chromosomes of humans were visualized and enumerated, and the chromosomal basis of one form of cancer (chronic myeloid leukemia) was identified. Following closely on the demonstration of polymorphic forms of hemoglobin S and C by electrophoresis, application of this technique to plasma proteins showed that protein polymorphism was a widespread phenomenon that merited further study.

New technologies that became available in the 1950s, combined with a more genetic approach to investigation, disclosed new relationships between the genetic control of responses to exogenous substances and their metabolic fate. Landmark studies of unexpected responses to primaquine, isoniazid and succinylcholine (suxamethonium) were the first to establish a link between drug response and heredity. The confluence of pharmacology, genetics and biochemistry noted in Arno Motulsky's 1957 seminal paper,⁽⁵⁾ concerned with person-to-person differences in drug response that arise because of the unique genetic constitution of individuals, marked the true beginning of pharmacogenetics as an experimental science. In 1962, Werner Kalow published the first systematic account of the field, demonstrating the implications of heredity for responses of microorganisms, insects, and vertebrates, including humans, to environmental chemicals.⁽⁶⁾ Since then, the characteristics of many more human traits of pharmacogenetic interest, too many to chart here, have demonstrated beyond any doubt that heredity is a remarkable prognosticator of unexpected response to drugs and other exogenous substances.

1.3 The Scope of Pharmacogenetics

Hereditary differences in toxicity or efficacy that result from exposure to therapeutic agents are usually considered as being within the province of pharmacogenetics, but if the unexpected response results from exposure to other chemicals, or to physical, climatic or atmospheric agents, it may be classified as ecogenetic.⁽⁷⁾ Ecogenetics encompasses genetically conditioned responses to exogenous substances from any source and, as such, would include the pharmacogenetics of drugs used in medical practice. It should be noted, however, that, despite this

distinction, the literature does not always adhere to this convention. When it is of interest to assess the characteristics of a given response, the criteria to be met in designing and executing a proper study to determine the heritability of the response in twins and families, and the distribution and incidence of genotypes and phenotypes in larger populations, are, in fact, identical for pharmacogenetics and ecogenetics.

Most well-studied polymorphisms derive from investigations of therapeutic agents, because these agents can be administered safely to individuals of all ages in defined amounts, and their disposition patterns can be measured. Historically, the drug-metabolizing enzyme polymorphisms are foremost in pharmacogenetics, and aficionados recognize these traits as the cornerstone of the field.⁽⁸⁾ However, since medical drugs and other exogenous chemicals are biotransformed by the same enzymes, unexpected effects, as Garrod noted, may extend to food-stuffs as well as to sporadic disorders associated with exposure to chemicals under occupational and industrial conditions. The enzymes that metabolize exogenous substances may also metabolize hormones and other endogenous messengers that regulate cell signaling; consequently, mutations that affect those pathways may also cause differences in individual response. This evidence serves as an additional hallmark of pharmacogenetic and ecogenetic phenomena. The range of such phenomena is virtually without limit.

The administration of two or more drugs in combination may also set the stage for an unexpected drug response that depends on the genetic makeup of the recipient. The CYP2D6 (debrisoquine/sparteine oxidation) polymorphism is a particularly rich potential source of such interactions. This polymorphism affects the disposition of more than 30 therapeutic agents including β -adrenergic blockers, antidepressants, antiarrhythmics, neuroleptics, and various miscellaneous drugs such as phenformin, dextromethorphan, and codeine. Combined administration of two or more of these agents may result in an aversive interaction.

Smoking represents another environmental hazard that results in a variety of responses and disorders which may depend on genetic predisposition of smokers. In individuals with a genetic defect in α_1 -antitrypsin, smoking is the most important environmental factor affecting the rate of deterioration of lung function.⁽⁹⁾ Additionally, information is accumulating that a variant form or forms of CYP2A6 may exhibit impaired capacity to metabolize nicotine and may thereby protect against the effects of smoking by reducing the number of cigarettes smoked.^(10,11)

Genetic predisposition appears to be involved as a hypersusceptibility factor for workers exposed to complex mixtures of chemicals. The formation of

aniline-hemoglobin adducts and excessive levels of methemoglobin can occur in those who are genetically “slow acetylators” after prolonged exposures to aniline derivatives.⁽¹²⁾ The sporadic occurrence of urinary bladder cancer in slow acetylators many years after lengthy exposure to the aromatic amine carcinogens⁽¹³⁾ employed in the manufacture of dyestuffs, plastics and electrical materials is a case in point.

Important clues to the role of heredity in unexpected responses to drugs and other chemicals are provided by individuals of different ethnogeographic origin. Perhaps the most familiar example is “primaquine sensitivity” occurring as a result of glucose-6-phosphate dehydrogenase (G6PD) deficiency among African, Mediterranean and Oriental persons. This is associated with sudden hemolysis, which occurs among susceptible males after exposure to any of some 200 drugs.⁽¹⁴⁾ Other examples of ethnic specificity in drug response include the protective effect against alcohol-induced liver disease of a variant of aldehyde dehydrogenase (ALDH2) that is found in Japanese but not Caucasians⁽¹⁵⁾ and the protective effect against alcohol-related birth defects of a variant form of alcohol dehydrogenase (ADH-3) occurring among African-Americans.⁽¹⁶⁾

While many of the examples cited above indicate that unexpected responses to exogenous chemicals may be mainly caused by a single (monogenic) factor, experience shows that variation of more than one hereditary element may contribute to the outcome. One example to illustrate the synergism that can occur between interacting factors is provided by a study of liver cancer caused by exposure to aflatoxin and hepatitis B among the Chinese. Aflatoxin exposure alone enhances the relative risk by approximately twofold, while infection with hepatitis B virus alone enhances the risk by approximately fivefold; however, combined exposure yields a 60-fold increase in relative risk of liver cancer. Hepatitis B infection, therefore, enhances the carcinogenic response among Chinese to aflatoxin by 30-fold.⁽¹⁷⁾

Nowadays, the scope for inherent variation in human response to exogenous chemicals is so great, and the pace of science so rapid, that it is difficult to stay abreast of new developments in the field. Although metabolic polymorphisms of drug-metabolizing enzymes appear most frequently in the literature, polymorphisms of other genes of pharmacogenetic interest, such as those that encode receptors, clotting factors, transporters and immunologic response proteins, appear increasingly frequently. In the 1990s, recombinant DNA technology has contributed immensely to our knowledge of pharmacogenetics by facilitating the development of sensitive and specific diagnostic methods for collecting molecular information about unexpected responses to drugs and other exogenous substances.

1.4 The Pharmacological Profile of Human Drug Response

Responses to a particular drug actually result from the physiological and biochemical attributes of cells that carry the necessary receptors in tissues of the recipient, but these responses are subject to modification by a variety of intrinsic and extrinsic influences.⁽⁴⁾ The age, gender, weight and other physiological and pathophysiological attributes of the individual, including heredity, are important intrinsic factors that can modify drug responses in individuals. Foods and other dietary components, exogenous chemicals in household and workplace environments, the use of tobacco and alcohol, and drugs (prescription and nonprescription) themselves are some of the extrinsic factors that may affect these responses.

The response of an individual following administration of a drug (or after exposure to an exogenous chemical) is a complex process that is difficult to analyze. By reducing this process to pharmacokinetic and pharmacodynamic mechanisms, pharmacologists have devised a somewhat idealized picture of the response, which can be analyzed more readily. Pharmacokinetic mechanisms are those that affect the concentration of the drug at receptor sites and its time course of action (absorption, distribution, and elimination by metabolism and excretion) in the individual, while pharmacodynamic mechanisms refer to receptor-mediated and allied events. This maneuver also provides a framework well suited to discussion of the pharmacological nuances of variation in human drug response.

The response of an individual to most drugs can be thought of as a train of events that begins when the drug enters the bloodstream and ceases when the drug and its metabolites are completely eliminated. The administration of a drug, given in ordinary doses, is usually accompanied by a gradual rise in drug concentration in the blood that reaches a peak or steady state; if no more drug is given, the concentration steadily declines until the drug and its metabolites are eliminated. If the drug attains a concentration within the therapeutic range, the characteristic response is expected. If the concentration reaches a level above or below that range, either a toxic response or the absence of the expected response may occur. For unexpected responses, experience indicates that there are two likely explanations. If the drug concentration in plasma is above or below the therapeutic range, one would anticipate a pharmacokinetic mechanism is responsible for the unexpected response, but if the concentration is within the therapeutic range, a pharmacodynamic defect is a more likely explanation.

Consequently, from the pharmacological viewpoint, unexpected drug responses to drugs and other exogenous

substances may be regarded to a first approximation as peculiarities resulting from a defect in either the pharmacokinetics or the pharmacodynamics of the drug.

1.5 The Genetic Profile of Human Drug Response

Geneticists, in contrast to pharmacologists, might consider a drug as an environmental agent on which individuals exert important differential effects; therefore, they might regard an unexpected drug response as a reflection of a genetic difference between normal and abnormal responders.⁽⁴⁾ Geneticists, as a rule, are more interested in looking for differences that discriminate individuals rather than focusing on universal similarities, and they would seek to explain an unexpected drug response by drawing on the gamut of information from the normal and affected responders as well as from their biological relatives. From the information obtained, they would attempt to determine the relative contributions of heredity and environment to the unexpected response.

Genetic defects that cause unexpected drug responses occur sufficiently often in many populations to enable members to be divided into two or more relatively common types of responder. Hereditary variation in which such sharply distinct qualities coexist in a given population is referred to as genetic polymorphism. This term is usually used in genetics to refer to genetic loci for which variants occur with a frequency of 1–2% or greater, but in pharmacogenetics the definition is better based on phenotype. This issue has been the source of some discussion^(18,19) and is considered further below (section 2.3). To avoid contradictions between pharmacogenetic usage and genetic theory and concepts, this article will refer to the polymorphisms linked to human responses to drugs and other exogenous substances as “pharmacogenetic polymorphisms”. Pharmacogenetic polymorphism may include variation produced by chromosomal aberrations, which are detectable by cytogenetic techniques, but most result from smaller, genic lesions.

To assess the contribution of heredity to pharmacogenetic polymorphism, unexpected drug responses are explored at all levels of gene action, from the gene molecule itself, to the individual and their biological families, to populations of individuals. Broadly speaking, hereditary differences in drug response are characterized by two types of information: the genetics that characterize the polymorphic phenotypes and their molecular basis. The latter includes a description of the genetic heterogeneity at the level of DNA and of the protein variants that explain the trait. By studying a given trait from several points of view, the relative influence of human ecology and heredity and the mechanisms by which they occur are ascertained. For these purposes, the concepts

and techniques of human genetics, biochemistry, population and molecular genetics, and epidemiology are used in concert to analyze pharmacological, toxicological and epidemiological observations of individuals, twins, families and larger populations.

1.6 The Rationale and Aims of Pharmacogenetics

Pharmacogenetic investigation often begins with an anecdotal clinical observation on one or a few individuals who have experienced an unexpected response to a specific drug, or with an epidemiological observation on a group of individuals who have developed a disorder associated with occupational exposure to a chemical or complex mixture of chemicals.⁽⁴⁾ The main goal of pharmacogenetics, simply put, is to use knowledge gained in understanding the influence of heredity on human sensitivity (or resistance) to exogenous chemicals to avoid the occurrence of such responses in susceptible persons. Therefore, while the early detection of cancer of environmental origin, for example, may do little to change the outcome in an affected person, identification of the trait that led to its initiation may prevent its occurrence in others.

A number of practical problems can arise in the design and execution of pharmacogenetic studies. A complete assessment of the causes of a given trait may be exceedingly difficult to achieve, may require the collaboration of scientists with differing expertise or may require resources not readily available. For example, epidemiological evidence may suggest a disorder (say bladder cancer) is caused by a particular agent (or agents), but the agent or agents usually cannot be identified unequivocally solely by epidemiological evidence; clinical observation may suggest an adverse reaction is caused by a certain drug, but the physician who contemplates investigation of the reaction by further exposure of humans to the offending drug may face insurmountable ethical and methodological constraints; susceptibility to a toxicant can often be more thoroughly studied under highly controlled conditions in animals than in humans, but studies in animal models are slow and expensive, and animal responses do not necessarily translate closely to those in humans.

Establishment of definable end-points for the investigation of specific differences between normal and abnormal responders poses another problem. Consider the evaluation of drug treatment of hypertension as a case in point. In this instance, the definitive end-point would be measured by the difference in the incidence of strokes and cardiovascular disease between normal and abnormal drug responders, but this would probably occur years after the trait is recognized and modifying treatment undertaken. Another example of pharmacogenetic interest

concerns the evaluation of individual variation in susceptibility to cancer associated with prolonged exposure to an occupational carcinogen. The definitive end-point in this case is the difference in cancer incidence many years later between hypersusceptible and normal responders exposed to the carcinogen. It is apparent that accomplishment of trials to determine definitive end-points might require studies extending over many years involving many individuals and large populations, and the payoff is not guaranteed.

Investigations with more limited goals designed to measure intermediate or short-term changes in biochemical and molecular markers may lessen some of these difficulties. Biological markers that can detect early and subtle differences in individual response would be necessary for this purpose. The most suitable are markers that can be detected in small samples of tissue or body fluids that are accessible to sampling; they should be expressed differentially in accordance with differences observed in normal and abnormal responders and they should have a low probability of spontaneous change.

Several advantages accrue to study designs that focus on the use of intermediate- or short-term biomarkers. First, the marker could be used to document and characterize the heterogeneity in response of individuals to a given drug, and to determine its distribution in individuals, families, and other populations of interest. This would indicate the frequency of a given trait, its mode of inheritance, and its ethnic and geographic specificity. Second, the markers could be used to identify the genes and the mutations responsible for the trait in susceptible people. This information would reveal the molecular basis of the trait and would aid in elucidating disordered physiological mechanism(s) that account for the aberrant response. It would shed light on the mechanism of the response to the agent in normal (wild-type) responders. The marker would also be of further use in the development of noninvasive, inexpensive diagnostic tests to determine the phenotype and genotype of individuals for the trait. Third, for agents used in medical therapy, identification of the molecular basis of specific traits and the elucidation of their underlying mechanisms would rationalize the selection of specific agents for susceptible (or resistant) individuals and make it possible to tailor drug regimens to the individual patient. Fourth, since the exact cellular site that is defective must be known for rational design of therapy, the information obtained in the first three study parameters would help to guide discovery of new drugs and to improve the design and conduct of their clinical trial. Finally, studies with markers could have wider implications for patient welfare by alerting physicians, clinical scientists, and others engaged in patient care to the importance of biomonitoring of individuals who suffer sporadic illnesses while ingesting

certain drugs and of individuals who become ill after prolonged exposure to toxic chemicals associated with a specific occupation or a particular environment.

An important goal of studies with intermediate- and short-term biomarkers is validation of the markers by demonstrating their correlation with definitive end-points. It is reasonable to expect that these biomarkers will enable differences in human susceptibilities to exogenous chemicals to be detected, or suspected, at an early stage of exposure, and the liability of susceptible persons to be predicted long before the definitive end-point is reached. It follows that studies using such biomarkers could be accomplished in a much shorter time and at a much lower cost than those aimed at determining definitive end-points.

1.7 Testing Pharmacogenetic Hypotheses

Exploring the pharmacogenetics of unexpected drug responses can be a formidable task. No hard and fast rules exist to guide the investigator's advance planning, but there are some general principles that are widely applicable.⁽⁴⁾ Experience indicates that the response of an individual to a drug and its inherent variation are best considered in light of the pharmacokinetic and pharmacodynamic mechanisms that define the disposition and actions of drugs in individuals (see section 1.4). Insights into the relative importance of the pharmacokinetic and pharmacodynamic phases of the response to the overall variability in the response to a given drug in a heterogeneous group of drug responders can often be gained by a preliminary pharmacokinetic analysis.

In general, an experimental approach is almost inevitably based on testing a main hypothesis and one or more auxiliary hypotheses.⁽²⁰⁾ The novice may, in fact, proceed with only a vague notion of the hypothesis being tested, but skilled investigators know that it is advantageous to have thought about it explicitly. Auxiliary hypotheses are set forth as reasonable alternatives in case the main hypothesis proves false. Investigators should also remember that, for every discovery made with innovative approaches, many important advances are made simply by recognizing that an established method or familiar technique devised for one application can be used for another. In some cases, it is possible for the investigator to design a single experiment that decides the fate of a given hypothesis.

1.8 Summary

Individual differences in human responses to therapeutic agents and other exogenous chemicals often result from some derangement of the genetic material that may be transmitted from one generation to the next. The scientific

study of the effects of heredity on drug response is the province of pharmacogenetics. Its main purpose is to explain the toxicological effects of the derangements by rigorous characterization of their biochemical and molecular basis. In this way, it is hoped to understand the causes of human susceptibility to exogenous substances and avoid the occurrence of unexpected drug responses in susceptible persons.

2 THE CHANGING SCENE

2.1 Human Protein Diversity

Proteins are vital to virtually every process of biological importance because of their capacity to bind other proteins and small molecules with high sensitivity and specificity.⁽²¹⁾ The assortment of proteins that resides in cells and tissues requires that the synthesis and targeting of these molecules is carried out with a high level of fidelity according to the well-defined set of instructions encoded in the genes. The properties of distinct proteins ultimately define the link that exists between the genetic constitution of a given individual and his or her response to drugs. It follows that a structural alteration in a given gene, or an error in carrying out the program of instructions encoded in the gene, might reasonably be expected to change the structure or the amount of the protein synthesized, or its final destination. Briefly, mutation could cause an unexpected drug response.

The occurrence of G6PD deficiency, succinylcholine sensitivity and isoniazid acetylation polymorphism, detected in the 1950s, was regarded as a serious threat to drug usage in medical therapy, requiring further investigation. Examination of the pharmacological defect and molecular genetic basis of unexpected drug responses set the agenda for subsequent research, and studies quickly showed that changes in protein structure resulting from genetic change altered human drug responsiveness wherever it was sought, among proteins with widely disparate functions, in health and disease.

2.2 Some Biochemical Generalizations

By the 1970s, Harris and colleagues had sought to examine the genetic causes of human protein diversity from the biochemical, chromatographic and electrophoretic characteristics of proteins associated with certain single gene disorders.⁽²²⁾ Certain hemoglobinopathies and enzymopathies provided the best models for this purpose because information about structural gene mutations was matched by physiological and pathological data. Evidence suggested that mutations might range from a single base change within the coding region of a gene to large deletions that removed a large part, or all, of a gene. A

single polypeptide chain that was affected by different mutations in the coding region might thus be represented by several distinct alleles. Most hemoglobinopathies, for example, could be ascribed to missense, nonsense, or frameshift mutations of the coding region of globin genes. Among the enzymopathies, observations suggested that allozymes (i.e. enzymes encoded by allelic loci) generated by point mutations and small deletions would usually display highly homologous amino acid structures and might have similar enzymatic properties, whereas larger deletions resulted in drastically shortened, nonfunctional gene products, or the total absence of the product.

Post-translational modifications were recognized as additional sources of the structural modification of proteins.⁽²²⁾ Should such a modification occur by an enzyme-mediated process, as had been established for oxidation of sulfhydryl groups or the addition of carbohydrate or phosphate groups, or by the cleavage of the polypeptide with loss of a terminal amino group, or a larger part of the chain, it too could be subject to genetic variation.

Unfortunately, the available biochemical data were inadequate to yield a detailed picture of the molecular basis of protein heterogeneity. For instance, a change that alters the charge on a protein can be detected by electrophoresis, but since two-thirds of amino acid substitutions are neutral, they will not change the net charge on the protein and will not be detected in this way. As a consequence of the inability of the standard methods of protein characterization to distinguish variation owing to genetic variation at the level of translation from that which occurs post-translationally, or nongenetically, information regarding the genetic basis of protein diversity was very meager.

2.3 DNA Polymorphism: Some Practical Matters

The analysis of genomic DNA reveals that the human genome is highly polymorphic. Sites that exhibit alternative sequences at a particular chromosomal site can be used as genetic markers for the site, or for the chromosome bearing the site. Three types of polymorphic site are useful for identifying individuals who are predisposed to unexpected drug responses: those within genes that determine drug response, those within short variable DNA repetitive sequences (known as variable number tandem repeat (VNTR)) and those within microsatellites. When such a site is associated with a given trait and tracked within families and larger populations, it can provide information about inheritance patterns and the prevalence of the trait. Polymorphic sites can be studied by restriction fragment (Southern) analysis to survey chromosomal sites for polymorphism, or by DNA sequencing to determine the precise location of base changes that

define a polymorphism. Southern analysis cannot resolve polymorphisms that differ by only one or a few bases, and neither can it identify the polymorphic base, whereas the size and boundaries of the polymorphic site can be precisely determined with the polymerase chain reaction (PCR) used in DNA sequencing. This second approach is essentially a refinement of the first and is preferred in many applications.

Since the development of molecular techniques, remarkable progress has been made toward identification of genes responsible for pharmacogenetic polymorphisms and toward understanding the molecular basis of quite a few of these polymorphisms at the DNA level. Most of the progress has related to traits attributed to a single polymorphic gene, many of which have been cloned and sequenced.⁽²³⁾ For some, the gene of interest has been expressed in heterologous expression systems in quantities sufficient for biochemical and pharmacological characterization.

The VNTRs or minisatellites of the human genome may be repeated 100 times or more in different persons. Many VNTRs, numbering in the thousands, are well characterized. Restriction fragment analysis will produce different size fragments proportional to the number of repeats in the VNTR. Where the identification of traits has been slowed for lack of a sufficient number of suitable genetic markers, the use of VNTRs should alleviate this constraint.⁽²⁴⁾

A number of genes have been found to incorporate microsatellites within them.⁽²⁵⁾ Microsatellites are stretches of repetitive DNA sequences; if mutated, they are capable of disrupting cell function. Such mutations may serve as hereditary markers of genomic instability that may increase individual susceptibility to certain cancers. For example, the loss of microsatellite DNA from the androgen receptor gene has been associated with prostate cancer. This event is of pharmacogenetic interest because a paradoxical response to antiandrogen therapy (i.e. stimulation of tumor growth instead of the inhibition expected) was exhibited by a patient.⁽²⁶⁾

The extent of protein diversity in natural populations is immense, but the abundance of DNA polymorphism in the human genome appears even greater. Analysis of genomic DNA reveals the presence of a large number of these polymorphisms. On average about 1 in 500 nucleotides differs between two randomly chosen alleles. Only about 5–10% of these polymorphisms is detected by restriction analysis, and frequently these polymorphisms are unrelated to a clinical phenotype. This phenomenon has been demonstrated for the CYP2D6 polymorphism by Skoda and colleagues,⁽²⁷⁾ who examined genomic DNA from 53 persons using a total of 20 restriction enzymes to fragment the DNA in different ways. Tests of 13 enzymes yielded 14 allele-specific polymorphic patterns

at the *CYP2D6* locus, while the other seven enzymes revealed none. This study and others (see references in Skoda et al.⁽²⁷⁾) show that DNA polymorphisms are often functionally silent at the protein level, having no detectable phenotypic effect.

It follows that information about a given genotype is insufficient for unambiguous identification unless combined with information about the associated phenotype. Further evidence pointing to the importance of identification of phenotypes arises from numerous sources. These include studies of the potential dissociation of phenotype from genotype observed in certain races^(28,29) and in certain families,⁽³⁰⁾ the lack of resolution of heterozygous and homozygous subjects by standard metabolic phenotyping procedures,⁽³⁰⁾ the different phenotypes for different substrates for a given polymorphism,^(31,32) structure–function relationships,⁽³³⁾ gene–gene interactions,⁽³⁴⁾ and the patterns of disease⁽³⁵⁾ associated with polymorphisms attributed to genes with closely similar sequences.

We should also point out that since the vast majority of pharmacogenetic traits, unlike most genetically determined human disorders, usually cause no recognizable effect on the health of predisposed individuals, the importance of phenotype takes on added significance. Only when susceptible persons suffer an unexpected response from exposure to certain drugs, dietary constituents or other environmental agents is their predisposition revealed. Fortunately, many of the polymorphisms associated with genetically variable human drug-metabolizing enzymes, can be detected in advance with the aid of a number of test probes that are suitable for human phenotyping by in vivo or ex vivo testing procedures. A partial list of these probes is presented in Table 1.

For the reasons stated above, pharmacogenetic polymorphisms must be defined on phenotypic grounds. Any comprehensive description of a given pharmacogenetic polymorphism should include information about the phenotype that characterizes the polymorphism and the genotype(s) that explain the polymorphism.

2.4 Molecular Heterogeneity of Pharmacogenetic Interest

Recombinant DNA techniques rapidly superseded standard methods of protein characterization for genetic analysis during the 1980s, and reports of the molecular characteristics of numerous enzymes, including the P450 (CYP450) enzymes of drug metabolism, soon appeared in the literature. Because of their specificity for endogenous steroids, fatty acids, and prostaglandins, as well as for many drugs, environmental pollutants, and carcinogens, the CYP450 enzymes had already attracted a great deal of

Table 1 Test probe drugs for human phenotyping of human drug-metabolizing enzyme polymorphisms

	Test probe	Refs.
Phase I enzymes		
Aldehyde dehydrogenase	Acetaldehyde	36
Alcohol dehydrogenase	Ethanol	16
CYP1A2	Caffeine	8
CYP2A6	Nicotine, coumarin	11, 37
CYP2C9	Warfarin	8, 37
CYP2C19	Mephenytoin, omeprazole	8
CYP2D6	Dextromethorphan, debrisoquine, sparteine	8
CYP2E1	Chloroxazone, caffeine	8
CYP3A4	Erythromycin	8
CYP3A5	Midazolam	8
Serum cholinesterase	Benzoylcholine, butrylcholine	38
Paraoxonase/arylesterase	Paraoxon	39, 40
Phase II enzymes		
Acetyltransferase (NAT1)	<i>Para</i> -aminosalicylic acid	41
Acetyltransferase (NAT2)	Isoniazid, sulfamethazine, caffeine	42
Dihydropyrimidine dehydrogenase	5-Fluorouracil	43
Glutathione transferase (GST-M1)	<i>Trans</i> -stilbene oxide	44, 45
Thiomethyltransferase	2-Mercaptoethanol, D-penicillamine, captopril	46
Thiopurine methyltransferase	6-Mercaptopurine, 6-thioguanine, 8-azathioprine	47
UDP-glucuronosyl transferase (UGT1A)	Bilirubin	48, 49
UDP-glucuronosyl transferase (UGT2B7)	Oxazepam, ketoprofen, estradiol, morphine	50, 51

attention among biochemical pharmacologists. Additionally, metabolic studies conducted during the mid-1970s with sparteine or debrisoquine suggested a polymorphic variant of CYP450 might be responsible for the poor metabolism and unexpected responses of susceptible persons. Other studies with mephenytoin suggested a variant of a distinctly different CYP450 appeared to explain the poor metabolism and unexpected responses of susceptible persons to this drug. In 1982, the complete primary structure of a mammalian CYP450 was reported from recombinant DNA studies, and this breakthrough prompted a great deal of further investigation.⁽⁵²⁾ A recent count shows that more than 150 isoforms of CYP450 have been characterized as products of different genes, and at least 30 different human CYP450 enzymes have been purified, cloned, sequenced and characterized. Of these, less than ten (namely 1A1/1A2, 2A6, 2C19, 2D6, 2E6 and 3A4) are responsible for oxidation of most drugs and other substrates in the human environment.⁽⁸⁾

A chronology of selected human pharmacogenetic polymorphisms that have been identified since the emergence of pharmacogenetics in the 1950s is given in Table 2. It includes a cross-section of polymorphisms: several drug-metabolizing enzyme polymorphisms, several receptor polymorphisms (e.g. malignant hyperthermia, vasopressin resistance, insulin receptor resistance) and polymorphisms of some other protein variants of pharmacogenetic interest. The table also draws attention to the time that elapsed between the initial description of

the inheritance of a given polymorphism and the explanation of the polymorphism at the DNA level.

As Table 2 shows, molecular genetic studies identified some of the molecular defects in CYP2D6* associated with poor metabolism of sparteine and debrisoquine some 11 years after the trait was discovered. CYP2D6* (sparteine/debrisoquine oxidation) polymorphism has stimulated an enormous amount of investigation and we now know that this polymorphism results in three separable phenotypes: poor metabolizers, extensive metabolizers, and ultrarapid metabolizers. The CYP2D6* poor metabolizers have an impaired capacity to metabolize more than 30 drugs and are homozygous for an inactive or deficient CYP2D6 enzyme that has been modified by truncation or missense mutations of the gene *CYP2D6**. In contrast, ultrarapid CYP2D6* metabolizers possess an enhanced capacity to metabolize these drugs because they possess more than one copy of an amplified *CYP2D6**.

Pharmacogenetic polymorphism is well established for the genes of several other CYP450 enzymes including 1A1, 2C9, and 2C19, and is rapidly accumulating for 1A2, 2A6, 2E1, and 3A4.⁽⁸⁾ Most of the genetically variant CYP450 enzymes occur as high- and low-activity (or null) isoforms that may confer individual susceptibility to the toxic effects of environmental chemicals including carcinogens. Molecular genetic studies have identified the genes that encode these enzymes as well as numerous allelic forms of them.

Table 2 also lists several drug-conjugating enzymes. During the 1990s, considerable progress has also been

Table 2 Chronology of pharmacogenetics

Polymorphism	Inheritance described	Mutation described	Elapsed time (years)
Succinylcholine sensitivity	1957–60	1990–92	33
G6PD deficiency	1958	1988	30
Long QT syndrome	1957–60	1991–97	34
Acetylation	1959–60	1989–93	30
Malignant hyperthermia	1960–62	1991–97	31
Glucuronosyl transferase	1966–69	1992–?	26
Vasopressin resistance	1969	1992	23
Aldehyde dehydrogenase	1969	1988	16
Debrisoquine oxidation	1977	1988–93	11
Retinoic acid resistance	1970	1991–93	21
Thiopurine methyltransferase	1980	1995	15
Mephenytoin oxidation	1984	1993–94	9
Glutathione transferase	1986	1990	4
Fructose intolerance	1986	1988–95	2
Insulin receptor resistance	1988	1988–93	0
Androgen resistance	1990	1990	0
Glucocorticoid remediable aldosteronism	1992	1992	0

made in the molecular genetic analysis of these enzymes. Prominent members of this group are the polymorphic forms of acetyltransferases, glucuronosyltransferases, glutathione-S-transferases and thiopurine methyltransferase.

Receptor proteins represent another subset of proteins (see Table 2) with enormous potential to influence human sensitivity to exogenous substances.⁽⁴⁾ Comparatively few receptor polymorphisms have been studied so far, since this group of proteins has not yet been explored in the depth achieved with the drug-metabolizing enzymes. Prior to the 1980s, the existence of receptors and receptor subtypes was customarily defined by differences in their interactions with drugs, but this approach did not achieve outstanding success in establishing receptor heterogeneity on a molecular plane. Based on their location and mechanistic features, two main types of pharmacological receptor are recognized: those that are located in the cytoplasm and with a locus of action within the cell nucleus, and those that insert into the cell surface and span the membrane. Cell surface receptors are involved in the actions of most drugs and endogenous first messengers; being on the cell surface, they can bind biogenic amines, protein and polypeptide hormones, autocoids, neurotransmitters and environmental chemicals. The nuclear (cytoplasmic) receptors, by comparison, interact with relatively few, albeit important, first messengers that can enter the cell, such as the steroid hormones. The techniques of molecular biology have enabled sufficient amounts of a given protein to be produced for biochemical and pharmacological characterization and have greatly refined and extended the genetic analysis of receptor heterogeneity.

2.5 The Effect of Mutations on the Gene Product

Mutations have been observed to have two main effects on the gene product: they can cause the production of a structurally variant protein with altered functional properties or the production of a fully functional protein in altered amounts, usually reduced, although traits attributable to increased amounts of protein have been described.^(53,54) Mutations that cause the synthesis of structurally variant proteins usually occur within the coding region of the gene, whereas mutations located within the promoter and other regulatory sequences usually alter protein quantity. Most pharmacogenetic polymorphisms result from point mutations, or other small genic lesions that lead to a functional change in a given protein, or loss of all or most of a gene. As these polymorphisms occur at specific genetic loci and within specific regions of the gene, the ability to detect and determine the associated specific structural change(s) allows for diagnosis.

Knowledge of the biochemical, pharmacological or toxicological features associated with an unusual phenotype can sometimes lead to inferences about the identity of the gene and the properties of the allelic variants responsible for the phenotype. Many pharmacogenetic polymorphisms were characterized by this approach long before molecular genetic techniques were developed. For example, observation of low levels of urinary acetylated metabolites of isoniazid suggested that an acetyltransferase, or acetyltransferases, with impaired capacity to acetylate isoniazid accounted for the isoniazid-induced nerve damage seen in susceptible patients. Biochemical studies documented the presence of functionally impaired forms of this drug-conjugating enzyme, which led to the

identification of genetically slow acetylators as a separate class of drug responders in human populations. The properties of acetylating enzymes isolated from tissues of slow acetylators were found to be closely similar to those from normal (rapid) acetylators, which suggested that a small genic lesion, possibly a point mutation, or even a small deletion, might account for slow acetylation of isoniazid.⁽⁵⁵⁾

At the time of the early biochemical studies, the existence of acetyltransferase variants could only be inferred; however, with the aid of molecular genetic techniques, the genes for acetyltransferase have been cloned and sequenced, and expression studies have demonstrated the existence of such variants, affirming predictions made some 30 years earlier. We now know that humans possess two drug-metabolizing acetyltransferase loci, and that variation at the *NAT2** locus is responsible for isoniazid acetylation polymorphism. We also know that the coding region of this gene is the site of multiple mutations that confer slow acetylation ability on individuals. A total of nine single nucleotide changes have been identified within the coding region of *NAT2** that account for more than 95% of slow acetylator alleles extant in human populations. Since these mutations occur singly and in combinations of two and three, they yield more than 20 *NAT2** slow acetylator alleles.⁽⁵⁶⁾

The *NAT2** polymorphism is an example of polymorphisms within the coding region of the gene, but recently polymorphisms within regulatory regions have also been described. The genetically polymorphic form of *CYP2E1* with enhanced activity that is associated with ethanol ingestion and with obesity represents an example of a regulatory polymorphism.⁽⁵⁷⁾ *CYP2E1* is recognized as an important enzyme for the detoxification of ethanol and many other xenobiotics and some drugs, including chloroxazone. Chloroxazone is a test drug that is used to detect intersubject variation in *CYP2E1* activity by in vivo and ex vivo procedures. Toxicological studies suggest a possible role for this enzyme in alcohol-related disease among African-American women. Analysis of the upstream region of *CYP2E1* revealed a DNA insertion, 100 base pairs (bp) long. The presence of this insertion mutation significantly increases *CYP2E1* metabolic activity as measured by chloroxazone hydroxylation, but only among persons who are obese or have recently consumed ethanol, or both. The frequency of heterozygous carriers of the mutation is 31% in African-Americans and 6.9% among Caucasians. This polymorphism in the regulatory region not only enhances *CYP2E1* activity but is sufficiently common to affect susceptibility to *CYP2E1*-related disorders in at least two ethnic populations.

These examples comprise only a small part of the molecular genetic studies that have been conducted on pharmacogenetic polymorphisms. They afford insights

into the genetic basis of phenotypic differences in human drug response, but it is prudent to avoid attempting to classify the effect of the mutation on the gene product from incomplete information. For example, a trait that appears, on limited study, to result from a quantitative defect in protein synthesis, may, on further investigation, be found to be caused by a structurally abnormal protein that disintegrates because it is inherently unstable or undergoes rapid proteolysis in the cellular environment.

2.6 Summary

Proteins are the molecular links between the genetic constitution and the drug response of the individual. The presence of a biologically active protein molecule at its native site in cells and tissues requires the synthesis and targeting of the molecule to occur in accordance with the well-defined program of instructions encoded in the genes. An alteration in the gene or failure to carry out the instructions with complete fidelity could produce a variation in the structure, function or final destination of the protein, any of which might alter the profile of drug response in the individual. Despite the remarkable advances in the understanding of pharmacogenetic polymorphisms at the molecular level, the dissociation of the phenotype from the genotype that is observed under various circumstances requires that both the phenotype and the genotype be characterized for unambiguous identification of a given polymorphism.

3 GENETIC PROFILING USING DNA MICROARRAYS

3.1 The Origin of DNA Microarrays

DNA microarray technology has its origins in Edwin Southern's method described in his landmark article "Detection of specific sequences among DNA fragments separated by gel electrophoresis" published in 1975.⁽⁵⁸⁾ There, he described how to combine the specificity of DNA restriction endonucleases with gel electrophoresis and sequence-specific hybridization to probe genomic DNA for its sequence composition. Restriction enzymes were used to generate controlled, predictable genomic DNA fragmentation, while gel electrophoresis provided an efficient, economic approach to arraying the fragments by size. The arrayed fragments were transferred by "blotting" onto a membrane support with a high affinity for nucleic acids, affording ease of handling and analysis of the same fragmented sample with a variety of probes. Finally, the complementarity of DNA base pairing as it occurs during hybridization permitted access to specific sequence information in the genomic DNA sample when it was hybridized with probes of known sequence.

The simplicity, elegance, and economy of the Southern blot facilitated its adoption into many molecular biology laboratories in a minimum amount of time. It not only became a standard laboratory protocol but also quickly began evolving into a diverse collection of hybridization-based analytical techniques. In 1977, for example, Alwine⁽⁵⁹⁾ described the Northern blot version of Southern's electrophoresis: a hybridization probe method to analyze messenger RNA (mRNA) sequences that provided the link between genomic sequence and functional mRNA expression analysis. Alwine's technique was also quickly adopted. Improvements and innovations in Northern and Southern blotting continue to simplify, refine and generalize the analysis of the broad range of sample types studied today.

Continued development in the Southern blotting method has greatly enhanced understanding of genomic DNA structure and organization. For example, routine use of Southern blotting led directly to the realization that restriction enzymes do not yield an identical DNA fragmentation pattern from every genomic DNA sample. Polymorphic variability occurring in the primary sequence occasionally ablates or creates new restriction sites in a genome. Restriction digestion and hybridization probing these samples results in altered banding patterns, called restriction fragment length polymorphism (RFLP).⁽⁶⁰⁾ Alex Jeffreys responded to this observation by developing "DNA fingerprinting" assays based on the Southern blot method.⁽⁶¹⁾ DNA fingerprinting relies on measuring complex combinations of RFLP fragments to document the degree of variability or similarity among individual genomic DNA samples. For this purpose, Jeffreys developed probes that hybridize with a type of DNA locus he discovered during his study of the myoglobin gene. These sites are characterized by VNTRs called minisatellites.^(62,63) These repetitive sequences, found throughout the human genome, are termed hypervariable because they exhibit variable numbers of repeats both within a single locus and between loci. VNTRs are still the most common markers used both for genomic DNA mapping and individual identity genotyping. Routine Southern blotting has resulted in the development of a large collection of informative VNTR probes suitable for many analytical applications.

Hybridization analysis using immobilized DNA includes "dot" blots and "slot" blots. Dot and slot blots are named for the circular and slotted well shapes in the templates used to present test samples to a membrane surface. They eliminate the need for restriction digestion and electrophoretic resolution steps by depositing samples of DNA to be tested directly onto a hybridizable membrane surface. Dot blot methodology is faster and easier for hybridization screening than Southern blotting, especially when many samples are to be screened

simultaneously using specific sets of hybridization probes. They can accommodate DNA samples too small or too damaged to undergo the purification, restriction digestion and electrophoresis demanded for Southern blotting, but tend to be somewhat less sensitive and informative than Southern blots since they do not provide RFLP information and are more prone to nonspecific background hybridization.^(64,65)

The invention of PCR permitted the development of the "reverse dot blot" in which the probe collections, rather than the DNA test samples, become the immobilized species on the membrane surface. DNA to be analyzed is prepared by PCR amplification of the region of interest, labeled with a detectable marker and hybridized to the array of immobilized probes.⁽⁶⁶⁾ The reverse dot blot concept led to the idea of generating generic arrays of probe sets that would allow high-throughput screening applications, as in routine molecular diagnostics. Commercial reagent suppliers began to develop generic products in response to the demand for common genomic analyses. The basic components that enabled individual investigators to customize reverse dot blot probe sets for their own applications have also become commercially available to investigators and are widely used throughout the molecular biology community.^(67,68)

As the density of information derived from efforts to sequence, map and identify human genes increased, so did the demand for analytical tools capable of exploiting this information. DNA microarrays were developed in response to this demand. Southern⁽⁶⁹⁾ was the first to describe parallel, in situ oligonucleotide synthesis as a means of generating oligonucleotide probe arrays on solid supports for highly parallel hybridization analysis. Southern's method uses standard nucleotide synthetic reactions to synthesize the oligonucleotides. The reactions are carried out in a movable chamber, which provides a physical barrier between the reaction chamber and the intended synthesis area.

This method of generating microarrays has been modified in several different ways since it was introduced. In one, photochemistry and photolithography have been incorporated into the oligonucleotide synthesis.⁽⁷⁰⁾ This technique uses novel monomeric reagents with photolabile protecting groups that are activated by light exposure; this permits controlled, localized array synthesis to be achieved. In another approach, piezoelectric nozzles and ink jet heads have been modified to "print" DNA synthesis reagents directly onto substrates that support in situ oligonucleotide synthesis localized by surface chemistry.⁽⁷¹⁾

In a sequence of events reminiscent of the Southern blot evolution into the Northern blot method, arrays of immobilized complementary DNA (cDNA) and expressed sequence tags have emerged on the heels of

DNA oligonucleotide arrays to profile mRNA expression patterns. In these “spotted” arrays, purified cDNA clones or PCR products are micropipetted as an array on a substrate surface and immobilized by one of a variety of covalent or noncovalent methods.^(72–74)

DNA microarrays are distinguished from reverse dot blots by their relatively high probe density, miniature size and use of solid, nonporous supports. Solid supports offer many functional advantages for arrays, as they become integrated components in automated, high-throughput assay systems. Rigid supports, particularly those that are optically transparent and thermally conductive, are more practical than large, flexible membranes for interfacing with automated fluid delivery or printing equipment, which introduce analytical samples, detection reagents and washing solutions. They also best accommodate the automated scanning systems that are used to image the arrays.

3.2 DNA Microarray Capabilities and Limitations

DNA microarrays already occupy niches in nearly every area of nucleic acid analysis. Both comparative or quantitative measurements and direct sequence analysis can be achieved, and microarrays have been designed and fabricated for applications as diverse as genomic mapping,⁽⁷⁵⁾ clone library screening, gene expression profiling,⁽⁷⁶⁾ genotyping and reference-based sequence checking (resequencing).^(77,78) However, despite their universal appeal and versatile functionality, DNA microarrays are analytical tools and should be evaluated in that context. The probe sets that make up an array and their associated functional assays are designed differently depending on the intended application. Arrays provide only a single component of complex biological assays that depend on well-integrated methods for isolation, amplification and labeling of the target to attain high sensitivity and high-resolution detection.

Sequence information is obtained from microarrays by decoding the complementarity between known array probe sequences and the partners they capture from a solution of labeled, hybridized test targets. Genotyping to identify a given polymorphism in a gene or DNA fragment can be used to illustrate how an array provides sequence information. A successful genotyping assay depends on developing maximum discrimination at the single nucleotide level among sets of closely related probes. This aim is best accomplished by selecting probe sets that display maximal signal uniformity and intensity as perfectly complementary duplexes. A genotyping result is interpreted by comparing positive signals that are anticipated from these duplexes with those obtained from sets of closely related mismatched probes. This method provides a background reference that controls

for hybridization specificity and permits confident interpretation of the hybridization result. Because mismatched probes do not hybridize well under optimal assay conditions, positive signals that display high intensity against a low nonspecific background are readily perceived.

During the development and fabrication of an array, candidate probes can be evaluated for optimal performance by using them to genotype target DNA samples that have been genotyped by a separate reference method. Confidence in genotype assignment can be improved further by selecting probes to interrogate the target sequence on both strands of target DNA and then combining the information to make a composite base call. Such an assay is not finished until an algorithmic analysis of its output is shown to be concordant with that of the reference for the same sample set. Discrepancies between the two sets of results must be resolved and the source of any difference must be attributed to the candidate or reference probes.⁽⁷⁹⁾

Defining acceptable performance specifications for a genotyping assay requires that a number of considerations be factored together. Because array hybridization and, therefore, genotyping quality are not likely to be completely uniform across complex hybridization arrays, it is critical to establish acceptable limits for the range of reporter and background signals observed; because the occurrence of different polymorphisms may differ in a given sample, individual polymorphisms should be weighted by the frequency at which each phenotype occurs, to achieve reliable prediction of composite genotypes; and because methods of target amplification and labeling of reporter molecules are variable, they may contribute significantly to the range of quality and confidence that will be observed in the results obtained.

Hybridization efficiency and hybrid stability both increase with probe length, and DNA array probes used for expression profiling tend to be longer than those used for genotyping.⁽⁸⁰⁾ This permits maximum detection capability for sequences of low abundance in complex RNA mixtures. Mismatch discrimination decreases as probe length increases, but high-resolution sequence discrimination is not a high priority in most expression profiling studies. Quantification of RNA abundance depends on hybridization signal intensity being directly proportional to hybridization target abundance, and maintaining this relationship depends on having a well-designed target preparation strategy. Linear target amplification methods such as *in vitro* transcription and ligase chain reaction coupled with controlled label incorporation into the amplified target have been quite successful in meeting this requirement. Appropriate sets of controls for use as internal standards for calibration of the assay can readily be developed.^(76,81)

The provision of effective gene expression profiling requires comparative and quantitative measurements of RNA. The target sequences in a given sample of RNA will represent a large concentration range of various transcripts, and the goal is to quantify them as accurately as possible. Strategies for expression profiling to meet this objective must maximize detection sensitivity for unique, low-abundance transcripts in complex hybridization mixtures dominated by highly expressed transcripts.

Effective expression-monitoring probes demonstrate hybridization sensitivity over a wide range of target concentrations. Global optimization of hybridization performance for expression-profiling arrays is as challenging as for genotyping arrays. The large number of probes in an array, coupled with the innate variability of the RNA pools used in the hybridization, make it impractical to test concentration dependence and sequence specificity for each probe independently. Consequently, other broad-based performance optimization strategies have been pursued. Designs that have proven successful for expression profiling include multiple hybridization probes for each mRNA sequence of interest, and related mismatch probes to estimate the contribution of nonspecific interactions to each specific hybridization signal. Redundancy also permits probes to be tailored to detect 3' and 5' ends and specific exons in mRNAs. These probes may be useful to gauge the quality of RNA and cDNA target preparations and to dissect splice variants or other closely related members of gene families. It has been reported that from four to twenty probe pairs (one fully matched to a target sequence and one mismatch control in each pair) may define the limits of useful redundancy in oligonucleotide-based gene expression profiling arrays.⁽⁷⁶⁾

A second, more common type of DNA microarray for expression profiling uses cloned cDNA or PCR products amplified from cDNA rather than oligonucleotides as capture probes. These DNA products are collected as libraries and robotically deposited onto chemically prepared solid substrates as addressable hybridization arrays. The original method for making these arrays relied mainly on noncovalent associations between the support surface and the one or two kilobase lengths of deposited DNA to keep the clones immobilized during hybridization.^(72,82) However, this method has rapidly evolved to include a wide variety of substrate surfaces, coupling chemistries and methods of delivering the probe DNA accurately and reproducibly to the surface. Both commercial and private efforts to optimize these "spotted" arrays are yielding ever more uniform and reproducible arrays. One advantage over oligonucleotide arrays that is gained with immobilized cDNA is that any clone from a cDNA library, whether or not it has been sequenced, can serve as a probe to screen for complementary sequences in other

RNA sources. In contrast, some primary sequence must be known for each target to generate synthetic oligonucleotide probes for an array. Hence, these two types of expression-monitoring array actually provide complementary functions. The other obvious contrast is that cDNA, because of its length, is likely to hybridize not only with its intended target but also with other closely related sequences, such as those within gene families. Oligonucleotide probes are more easily customized to discriminate among related sequences and so may offer a more refined expression profile for individual related sequences.

Microarrays based on cDNA or oligonucleotides differ fundamentally in how their experimental outputs are analyzed. Typical cDNA arrays are hybridized with differentially labeled cDNA pools generated from two separate RNA sources. For example, RNA may be harvested from untreated cells grown under a standard set of conditions and the cDNA produced from this RNA pool may be labeled with one fluorescent dye. A second RNA sample is then harvested from the same cell type after it is treated with a chemical or grown under a different set of conditions, and the cDNA is labeled with a second dye. Once the two cDNAs can be distinguished by their labels, equal amounts are mixed and hybridized competitively to the same cDNA array. The ratio of one dye signal to the other at each probe will reflect relative differences in abundance between the two RNA samples for the gene represented.

In contrast, a typical oligonucleotide array is usually hybridized with a single RNA pool labeled using a single reporter molecule. A measure of the hybridization intensity at each probe is obtained using internal calibration standards. The success of this approach depends on having a way to normalize the signal at each probe for any nonspecific background contributing to the signal intensity. This is usually achieved, as mentioned previously, by pairing a closely related, mismatched probe for each specific probe in the array. In addition, use of a redundant pair of probes for each target helps to average out this variability.

As the amount of available information about expressed genes accumulates and the sequence of the entire human genome emerges, successive generations of DNA microarrays have become increasingly dense in probe sites. The result is a dramatic increase in information gathered from each profiling experiment, and greater insight into the complexities of cellular biology. This gain is accompanied by target preparation requirements of increasing complexity to access the information that an array offers. What the useful limit to hybridization complexity will be is not yet clear. Entire cDNA libraries and the entire yeast genome have been successfully hybridized to DNA microarrays to yield large volumes of information

from single experiments, but a method for whole genome typing of a target organism in a single hybridization is still out of reach. Currently, PCR or some alternative strategy is required to achieve informative human genotyping by microarray hybridization. Variability in performance potential among hybridization probes and nonspecific interactions on the array surface (both within targets and between targets and probes) limit full reconstruction of such a complex target from a single hybridization.

The measurements made using DNA microarrays are very difficult to optimize to a tightly specified standard because of the high density of information and its parallel nature. It is equally difficult to validate individual probes globally in microarray hybridization assays at every data point, since it is impractical to check every point by a reference assay method. Statistical approaches to solving this problem include checking a selected sample of output data by an independent assay and testing statistical variability within and among repeated assays. A second strategy to validate the results of microarray hybridization is based on the integration of sets of complementary hybridization controls directly into the array.

The process of establishing and validating DNA microarrays is further constrained by the fact that their functional use and performance evaluation must be in the context of a broader, integrated system. That is, peak array performance occurs when probe design has been tuned to the chosen target, the labeling method is efficient and reproducible, and the imaging system is sensitive, calibrated and well matched to the reporter molecule used. Finally, sensible output analysis depends on the specific assay application and the control elements built into the assay. Arrays, especially very high-density arrays, are most efficiently used with automated interfaces that integrate all assay steps from sample preparation, through hybridization, data acquisition and output analysis. Ultimately, full microarray assay automation should become available from sample input to output of the genotype or expression profile. Progress in microarray development has been impressive, but an equally impressive amount of work remains to be done before these methods will be fully robust clinical tools.

3.3 Predictive Testing with DNA Microarray Assays

Predictive testing based on DNA microarray profiling is not available at the present time, although it is widely promoted. True predictive pharmacogenetic testing relies on having broad individual genotyping capability and integrated rational therapeutic management strategies for disease guided by genetic predisposition and risk profiles. Reaching this goal will require more than just technical development. The pharmacokinetic and

pharmacodynamic end-points that are used to describe therapeutic indications and responses today need to be integrated into a coherent picture and translated into molecular end-points. Our fundamental knowledge of the physiological effects of therapeutic agents must also be reduced to a molecular level. Pharmacogenetic traits are remarkable prognosticators of individual responses to drugs, exogenous chemicals and xenobiotics, but our present level of understanding of the pharmacokinetic and pharmacodynamic effects of drug response at the molecular level is rudimentary.

Our current state of knowledge of human drug response stems largely from retrospective genetic analyses guided by phenotypic responses to specific drugs. As these observations are studied in greater depth, it is becoming clear that few, if any, phenotypic outcomes will be reliably predicted from simple genetic analyses at a single locus. It is also clear that even when complex genetic analysis becomes more sophisticated and widely available, predicted phenotypes will still represent a potential rather than a certain outcome. Translating predicted potential into an outcome prediction will still depend on factoring environmental variables such as diet, age, drug use, health status and environmental exposure history into the equation. It is also likely that clinical monitoring of therapeutic responses will not be replaced by, but rather will be guided by, predictive testing. For example, a patient may have a genotype that predicts metabolic tolerance for a cardiotoxic medication with a narrow therapeutic index. The drug may be given to him because it most effectively targets his genetically profiled therapeutic need. However, blood monitoring for toxic drug levels may still be required since the effect of illness on his metabolic performance is not predicted by his genetic profile.

Two radically different approaches have been proposed for developing therapeutics based on known molecular mechanisms for individuals whose phenotypic responses are projected from their individual genetic profiles. One approach is to generate a high-resolution map of the human genome using characterized polymorphic markers, especially single nucleotide polymorphisms. Once the map is available, large-scale population epidemiology and case-control studies are expected to provide high-quality associations between genetic markers, specific disease and therapeutic response phenotypes. These population-derived markers are anticipated to be useful for projecting individual outcomes and risk assessments. The alternative approach has been to find individual genes that give rise to distinctive phenotypes and study them in depth. Each gene is fully characterized for polymorphic variability and each variant is characterized by its impact on the phenotype. As this information is assembled, phenotypes may be recognized that cannot be explained by known genetic

variants, and genetic variants will be recognized that do not lead to separate phenotypes. These observations may, in time, lead to the discovery of new functional relationships with other genes. Both strategies claim some success and both contribute to our basic understanding of human genetic variability. Whether one approach is superior to the other remains to be seen.

3.4 Emerging Applications of Microarray Technologies

Microarray technologies are useful for many applications requiring direct nucleic acid sequence analysis, sequence comparisons or quantitative comparison among nucleic acid samples. Limitations exist where direct hybridization analysis is not the best-suited approach, as for analysis of trinucleotide expansion (fragile X and Huntington's disease) and microsatellites. Microarray applications are not limited to human genetics but can also be used for target organisms such as yeast, model organisms *Arabidopsis thaliana* and *Caenorhabditis elegans*, viral and bacterial human pathogens, mouse, rat and ape genomes, and a wide variety of other agricultural plant, pest and animal targets. It is likely that this entire spectrum of genomic studies will ultimately contribute significantly to clinical applications of pharmacogenetics. Human gene complementation applied to regulatory pathways in model organisms has a strong track record for predicting parallel human biology. In addition, these models provide powerful pharmacokinetic and pharmacodynamic discovery assays and toxicology screening systems. Continued expansion of its technical foundations is expected eventually to provide combined genomic profiling for genetic variability and functional expression profiling capability as single assays.

All early microarray assays relied on hybridizing a labeled target or test nucleic acid directly to probes covalently linked to a substrate surface in an addressable format. Signal location provided sequence information about the hybridization target; signal intensity provided semiquantitative information about target abundance in the hybridization solution. The need for improved quantitative and discriminative performance from microarray assays has led to rapid innovation and improvement in a number of different areas: in the design of initial arrays and assays; in modifications in nucleotide chemistry,⁽⁸³⁾ linker chemistry,⁽⁸⁴⁾ substrate preparation chemistry⁽⁷¹⁾ and hybridization solution compositions,⁽⁸⁵⁾ which permit ever-increasing refinement and control over assay performance; in improved array components, which permit performance of increasingly complex and sensitive analyses; and in the coupling of primary hybridization assays to secondary enzymatic processes, such as primer extension and enzymatic mismatch cleavage, to increase the quantity and quality of information yielded by a single assay.^(86,87)

Applications of microarrays to genotyping include typing of samples for known sequence variants and discovery of new polymorphisms in specific targets, such as drug-resistant pathogens and functionally variant forms of known genes. Arrays for these types of test perform best with probes that have similar hybridization behavior, and efforts to design such probe sets have proven successful. Hybridization to complete collections of oligonucleotide probes of a given length shows a wide range of variability that correlates directly with individual nucleotide composition. An alternative method of design that avoids such variability is to target array probe sets to a narrow range of thermal stability. Probes designed by this criterion may still vary in length, but they show much more uniform behavior within a single set of hybridization conditions. Discrimination of the hybridization signals of fully complementary hybrids from the uninformative signals that result from associations between partially matched nucleotide fragments is much improved by this approach. Further refinement of hybridization uniformity has been obtained by incorporating nucleoside analogs that smooth thermal stability and disrupt intra-probe hairpin structures and mismatch associations. These assay modifications help to maximize discrimination among specific sets of related nucleotide sequences. These developments have been important in the effort to provide arrays for sequencing applications where samples are hybridized to an array of oligonucleotides that reconstructs a consensus sequence.^(83,88) Probes that fail to hybridize to this type of array indicate the presence of a polymorphism in the test target sequence that causes loss of complementarity with the consensus probe.

The density of information from microarrays has increased in parallel with technical advances in robotic microfluidic delivery systems and photolithographic techniques for fabrication of arrays. The result is an increased yield of information from single assays, but this comes at the cost of the increased complexity in preparing a test sample for hybridization in a way that uses the array optimally. For genotyping or sequencing applications, this usually means that a large number of PCR products must be generated in essentially equal quantities and labeled as uniformly as possible with a reporter molecule that permits hybrid detection. For expression-profiling experiments, RNA must be quantitatively isolated from the test source, amplified enough to permit rare mRNA detection but not enough to lose the quantitative relationships among the various mRNA species. Uniform labeling of samples with a reporter molecule is critical to keep the hybridization signal of each mRNA proportional to its abundance. These analytical demands have been met with a variety of innovations in array design and target labeling strategies.

As microarray technology has advanced, an abundance of new and challenging analytical applications have arisen, for example cancer biology. Many cancers are characterized by nucleic acid changes in somatic cells. The challenge here is in finding efficient ways to present the very limited amounts of test sample available from a biopsy to arrays for hybridization analysis. Loss of heterozygosity is another important somatic cell change indicative of cancer progression that can be detected only by quantitative genetic analysis. Detection of drug-resistant pathogens is another quantitative analytical problem, since genotypes of resistant organisms emerge against a background of susceptible genomes. Similarly, assessment of host selectivity for a given pathogen demands the detection of virulent subtypes occurring at low frequency in a background of benign genotypes. Studies to document human genetic variability for association studies to link our genes to our environmental responses on a population basis require true high throughput genotyping for large numbers of genetic markers. These studies present challenges in the technical aspects of genotyping by microarrays and in bioinformatics.

3.5 The Future of Microarrays and Allied Technology

Despite remarkable innovation and rapid technical improvement in microarray technology, several significant limitations in its application remain. Methods of fabrication need improvement to enhance their utility and reduce costs of construction. The availability, utility and efficiency of bioinformatics methods used to access, process and store array information are also limited. Preparation of an adequate quantity of high-quality, labeled target DNA for a given assay still presents a major difficulty in the use of DNA microarray technology. This is particularly problematic for gene expression profiling, where quantitative output is important and the relative abundance of the starting materials must be preserved during amplification and labeling. In addition, as more and more of the genome is accessed for profiling and genotyping, the problem of preparing amplified material from all relevant sites in the genome presents significant technical and economic burdens. At present, methods of multiplexed PCR amplification are not sufficiently robust or reliable to permit quantitative production of target material on a genomic scale. In addition, the cost of enzymatic amplification and labeling reagents for large-scale target preparation is significant.

A better method for highly parallel genetic analysis is needed. One with single molecule sensitivity that eliminates both the requirement for target amplification and the need for a target to be chemically modified or labeled for its detection would be ideal. Innovations in array construction, sample nucleic acid preparation

and detection hold this promise for a second generation of microarray technologies. For instance, Heller and colleagues have developed arrays with bioelectronic chips that are electronically addressable.⁽⁸⁹⁾ An electric current can be used to address each feature of the chip, enabling direct capture probe immobilization as well as control of hybridization stringency with the target molecules. This type of array has been shown vastly to accelerate target capture rate by using electric fields at each probe site to overcome the passive diffusion that previously had been a rate-limiting factor.^(90,91)

A number of completely different approaches to array-based genetic analysis show considerable promise. One consists of a DNA microarray combined with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI/TOFMS) detection. Technical development of this system is moving ahead rapidly. MALDI/TOFMS detects sequences by measuring their mass, which is attractive because it does not require separate labeling of the target. Initially, the principal limitation of MALDI/TOFMS lay in its inability to discriminate large molecules with sufficiently high resolution, but strategies have been devised that mitigate this problem. A second disadvantage of MALDI/TOFMS is the relatively large amount of sample material needed to achieve reliable results. MALDI/TOFMS systems are commercially available, and the development of methods to apply this system to determination of genotypes is accelerating.^(92,93)

Another unique approach to microarray technology uses fiber optic bundles to support an array of microbeads that are coated with DNA capture sequences. Each bead is individually associated with a single fiber in the bundle, and hybridization targets labeled with fluorescent dyes are hybridized to the beads. The one-to-one concordance of bead to fiber in the bundle permits every bead in the hybridization array to be addressed and reported individually.^(94,95)

DNA microarrays, even at their current state of technical evolution, are proven analytical tools with sufficient capacity to meet future needs of genetic assays. There are, however, technical challenges that must be met before they can reach their full potential, as well as significant challenges in bioinformatics and information management that applications of microarrays have created. Microarray technology generates such vast quantities of information that powerful databases, only now emerging, are necessary to analyze the information obtained and to store it in a format that is accessible. New concepts in database mining will be necessary to relate these large, complex sets of genetic data to the phenotypes recorded in clinical databases. The new knowledge gained from the

discovery of associations between genotypes and phenotypes of individuals in response to exogenous agents should enable clinical applications of pharmacogenetics to develop.

3.6 Summary

Currently, microarray applications span all aspects of molecular biological investigation, enabling the collection of large amounts of genetic data from comparatively few experiments. At present, microarrays are used mainly as research tools to determine the most effective ways of gathering information for individual genetic profiles and management of unexpected responses to medical therapy and environmental substances. With further development, they should contribute substantially to point-of-care analysis and therapeutic management. The full potential of the technique will only be realised when bioinformatics to analyze large datasets are also fully developed.

4 THE OUTLOOK FOR GENOMICS IN PREDICTIVE MEDICINE

4.1 Options for Treating Genetic Disorders

The care of patients with genetic disorders involves many of the approaches to diagnosis and treatment that are used in other medical specialities, but the focus in medical genetics is more toward prevention or avoidance of the disorder. Presymptomatic (including prenatal) diagnosis, genetic screening programs and genetic counseling are concepts central to this approach.

Options for treatment of genetic disorders at the environmental level routinely involve a combination of restriction, replacement and removal of the toxic substance. Restriction of potentially toxic environmental substances could, for example, involve restriction of certain foods or other dietary constituents. Fructose is one of the major constituents of human diets, and intolerance for this sugar is represented by a polymorphism in the aldolase B gene.^(96–98) Continued ingestion of fructose or its congeners is required to establish the disease, but the wide distribution of these substances in foods and some drugs places genetically susceptible persons at constant risk from an avoidable nutritional disorder. The fish odor syndrome (trimethylaminuria) is another example of variability in human response to foods.⁽⁹⁹⁾ A polymorphism in the gene for flavin-monoxygenase 3 accounts for some cases of this trait.⁽¹⁰⁰⁾ Trimethylamine, a metabolic product of numerous foods that contain choline or carnitine, confers the smell of rotting fish on affected persons, with devastating educational, economic

and social consequences. Management of trimethylaminuria remains empirical; the main recourse for affected persons is dietary restriction to reduce their intake of trimethylamine precursors. This may require avoidance of foods rich in choline, such as eggs, liver, soya beans, and marine fish.⁽⁴⁾

Restriction of certain therapeutic agents is an option that is commonly applied by physicians to avoid adverse drug reactions in genetically susceptible patients. Individuals with a deficiency in G6PD (Table 2) must avoid oxidant stresses such as those that accompany intake of antimalarials, sulfonamides and an extensive list of other drugs; failure can result in hemolysis of red blood cells. Individuals with thiopurine methyltransferase deficiency (Table 2) usually must receive much smaller doses of 6-mercaptopurine, 6-thioguanine, and 8-azathiopurine if they are to avoid toxicity (bone marrow suppression) by these agents. These drugs are mainstays of antileukemic and immunosuppressant therapy. Those with this inherited deficiency are intolerant to these agents and may suffer acute and delayed responses that can be life threatening.⁽⁴⁶⁾

Toxicity resulting from exposure to cigarette smoke in persons afflicted with α_1 -antitrypsin deficiency has been mentioned above (section 1.3). Avoiding cigarette smoke can prevent its destructive effects on the lungs, which result in emphysema and chronic obstructive lung disease.⁽⁹⁾

Replacement of deficient gene products or even of organs is also utilized in the treatment of genetic disorders: for example, replacement of coagulation factor VIII in hemophilia A, of α_1 -antitrypsin in persons deficient in this factor or of pancreatic islet cells in some forms of diabetes mellitus.

The option to remove the toxin from the environment of susceptible persons is taken in hemochromatosis, a common disorder of iron metabolism⁽¹⁰¹⁾ that affects 1 in 300 persons of northern European descent. Most cases of hereditary hemochromatosis are attributed to a polymorphism of *HFE*, which results in substitution of tyrosine for cysteine at codon 282. The disease, untreated, causes liver cirrhosis, heart failure, diabetes and arthritis and leads to early death; treatment by phlebotomy to remove excess iron allows affected persons to live a normal life span.

4.2 Many Pharmacogenetic Polymorphisms are Exceptions to Experience with Single Gene Disorders

The point was made above (section 1.6) that pharmacogenetic polymorphisms usually have no perceptible effect on the health of predisposed persons, unlike other genes involved in human diseases. The frequent occurrence

and ethnic specificity of human pharmacogenetic polymorphisms are further features that set them apart from many human disease genes.

The gravity of effects that may be in store for susceptible persons is only revealed by exposure to certain drugs, dietary constituents, or other environmental toxins; without this exposure, their genetic predisposition is invisible. Most people perceive heredity as having an important effect on health and disease, but may be unaware of the relevance of genetics to their own responses or of person-to-person differences in response to environmental toxins. The belief is widely held that unexpected responses to these substances are entirely linked to the toxin, but we know that the drug recipient is not a passive participant in these events.

Hemochromatosis (1/300)⁽¹⁰¹⁾ and cystic fibrosis (1/2000 to 1/2500)⁽¹⁰²⁾ are regarded as highly prevalent genetic disorders in human populations, but they are rare compared with many pharmacogenetic polymorphisms. Polymorphisms of *G6PD*, *CYP2D6*, *CYP2C19* and *NAT2* illustrate this point very well.⁽¹⁰³⁾ *G6PD* deficiency is a sex-linked trait that affects 5–10% of dark-skinned races. Males of African, Mediterranean, and Oriental descent are particularly susceptible. Globally, this deficiency is estimated to affect more than 400 million persons.

CYP2D6 polymorphism has been the subject of many population studies. More than 95% of deficient alleles responsible for *CYP2D6* poor metabolizers have been identified (see section 3.2). Eleven nucleotide changes belonging to seven *CYP2D6* allelic variants describe the extent of ethnic variation in people of Africa, Asia and Europe.^(104,105) The prevalence and ethnic specificity of the ultrarapid *CYP2D6* phenotype described recently indicates that Swedish and German populations possess 1–2% of this phenotype, whereas the frequency is higher in Spain (3.5–5%); the carrier (heterozygote) frequency in Ethiopia and Saudi Arabia is also very high (15–20%).⁽¹⁰⁶⁾

The prevalence and ethnic specificity of *CYP2C19*, another polymorphic P450 enzyme (Table 2), also varies quite remarkably. *CYP2C19* polymorphism, previously named mephenytoin polymorphism, is of major importance because of its role in elimination or activation of drugs such as omeprazole, an antiulcer drug, proguanil, an antimalarial drug, and several barbiturate hypnotics and sedatives.⁽¹⁰⁷⁾ The frequency of poor drug metabolizers owing to the m1 (*CYP2C19**2) variant is high in Japanese (13–23%) compared with Caucasians (4%). Another variant, m2 (*CYP2C19**3), occurs in Japanese and Africans but has not been detected in Caucasians. The wild-type and the m1 and m2 variants account for almost 100% of the variation in Asian poor metabolizers, but these and additional variants account for only about 92% of the variation in Caucasians.⁽¹⁰⁸⁾

Studies of the prevalence and global distribution of *NAT2** polymorphism have involved well over 10 000 subjects in dozens of populations since the polymorphism was identified in the 1950s (Table 2 and section 2.5). The percentage of slow acetylators ranges from 80% or more in Egyptians and certain middle Eastern populations to 20% or less in Japanese and Canadian Eskimos. Populations of European and African origin are, with few exceptions, characterized by intermediate percentages of the slow acetylator phenotype.⁽⁵⁵⁾ Molecular studies indicate that *NAT2** allelic frequencies are neither uniformly nor randomly distributed across different populations. For instance, the frequency of the rapid acetylator *NAT2**4 allele is approximately 20–25% among American and European Caucasians, 36% among African-Americans, 42% among Hispanics, and 66–70% among Asians from Hong Kong, Korea and Japan. The distribution of *NAT2** alleles for slow acetylation varies with racial origin; in Caucasians, three alleles (4, 5B and 6A) account for about 95% of all *NAT2** alleles, whereas in Oriental populations the 5B allele becomes rarer (5% in Hong Kong Chinese to <1% in Japanese) and 6A (20–30%) and 7A/7B (7–16%) are most frequent.⁽¹⁰³⁾

4.3 Emerging Applications of Pharmacogenetic Interest

Areas of focal importance for the pharmaceutical industry and for those engaged in the development of microarray and allied technologies include adverse drug reactions, diagnostics and targeted therapies. The versatility and broad capability of microarray technology permits it to be applied to any of these areas for the study of variation in genes, pathways and targets associated with disease.

The main challenge for the immediate future is to devise reliable methods and techniques for identifying and scoring all types of genetic variation in the human genome. Several strategies and technologies for identification of such variation at the molecular level are being developed (sections 3.3–3.5). Questions have been raised as to whether microarrays have the capability to identify different types of mutation, and whether they can distinguish heterozygotes. Cronin et al.⁽⁷⁹⁾ examined the first question in a study of *CFTR*. They found that the gene, composed of 27 exons, contained more than 500 widely distributed mutations, including transitions, transversions, insertions, deletions and polymorphisms. Two microchip arrays (480 and 1480 probes of 14–16-mers) were used to scan exon 11. Ten unknown genotypes assigned by these assays were identical with PCR product restriction fragment analysis.

Hacia et al.⁽⁷⁷⁾ looked at detection of heterozygotes predisposed to breast and ovarian cancer in a study of *BRCA1*. They found that it comprised 22 exons spanning 100 kb and contained more than 110 mutations

(frameshift, nonsense and missense) distributed throughout the gene. A microchip array with 96 500 20-mer nucleotides used to scan 3.45 kb of exon 11 accurately diagnosed 14 of 15 patient samples, and no false positives were found among 20 controls. Even though the full potential of these devices is not yet realized, these studies, among the first to report applications of microarrays to genes associated with disease, demonstrate very clearly the capability of this technology for genetic analysis.

The analysis of pathways affected in disease is another emergent application of this technology. Pathways that regulate cytokine signaling, insulin signaling and apoptosis are among those identified for early investigation. Targets of high priority for study include G-protein transmembrane receptors, ion channels and transcription factors. The nucleic acids targeted for these and some other emerging applications of focal interest are mRNA (cytokine induction, tumor suppression), DNA (cytokine mutants), germline DNA (*BRCA1*), yeast DNA (bar-coding mutants), DNA of human immunodeficiency virus (protease-resistant polymorphisms), mitochondrial DNA (natural polymorphisms) and *CYP2D6* and *CYP2C19* DNA (drug-metabolizing enzyme polymorphisms).

4.4 Linking Human Genotyping to Clinical Applications

The ultimate goal of pharmacogenetics is to gather information that explains unexpected responses to exogenous chemicals and that prevents or avoids the occurrence of these responses in genetically susceptible persons (see section 1.6). Among all the possibilities emerging from, or accelerated by, the human genome initiative,⁽¹⁰⁹⁾ pharmacogenetics is an area that may rapidly bring the predictive prospects of human genotyping into the clinical arena. In essence, a given pharmacogenetic polymorphism is characterized by three types of information. These are the genetics (mode of inheritance, allelic frequencies, and ethnic specificity), the molecular basis (genes responsible, and their mutation spectrum) and the medical or biological significance of the polymorphism.

Ethnic specificity not only provides information about unique features of a population that is of evolutionary interest, but may also explain situations of therapeutic interest. Consider the administration of a drug metabolized by the *CYP2D6* system, such as codeine, to patients of different ethnicity. Codeine is an analgesic prodrug, the effectiveness of which depends on its conversion to morphine. This conversion is impaired in poor *CYP2D6* metabolizers, and consequently they do not get pain relief from ingestion of codeine.⁽¹¹⁰⁾ For the same reason, they are also protected against dependence on oral opiates such as codeine.⁽¹¹¹⁾ The failure to respond to codeine would, consequently, be more prevalent among Africans and Caucasians (5–10%) than among Asians (<1%).

Unexpected neurotoxicity from amphetamine and its analogs (e.g. 3,4 methylenedioxymethamphetamine, also known as “ecstasy”) as well as interactions with other drugs subject to *CYP2D6* polymorphism would likewise be expected to occur more frequently among Africans and Caucasians than in Asians.⁽¹¹²⁾

The analysis of the molecular basis for unexpected responses of ultrarapid *CYP2D6* metabolizers to codeine follows a similar line of reasoning. The ultrarapid metabolizers have enhanced capacity to metabolize codeine and hence may exhibit exaggerated responses, such as abdominal cramping, fuzzy vision and disorientation.⁽¹¹³⁾ Because ultrarapid metabolizers occur more frequently among Hispanic, African and Saudi Arabian than Asian populations,⁽¹⁰⁶⁾ the former groups would be more likely to experience *exaggerated* responses to codeine.

4.5 Summary

Genetic disorders with abnormal reactions to environmental factors are no less treatable than many other disorders treated by physicians. These genetic disorders are treated by restricting access to, replacing and/or removing the toxic substance from the environment of the susceptible persons. Pharmacogenetic polymorphisms are invisible until their effects are revealed by exposure of individuals who harbor them to environmental toxins; microarrays and allied technology provide the means to screen for these polymorphisms and to identify susceptible individuals before they are affected.

5 SUMMARY: PROGRESS IN PHARMACOGENETIC TESTING

Individual susceptibility to the effects of drugs and other exogenous chemicals is often the consequence of some derangement of the genetic material that may be transmissible from one generation to the next. The scientific study of these effects is the province of pharmacogenetics. The main purpose of pharmacogenetics is to understand the causes of unexpected responses to exogenous substances and to prevent or avoid their occurrence in susceptible persons. Prior to the development of recombinant DNA technology, the role of changes in the genetic material and the consequent production of aberrant proteins, or failure to produce proteins, in disease could only be inferred. The development of microarrays and allied technologies were direct outgrowths of recombinant DNA technology and can be used to identify and screen for all types of molecular genetic variation affecting genes, pathways and molecules relevant to human genetic disease, including variations of pharmacogenetic importance. Microarray

applications currently span all aspects of molecular biological investigation, mainly in research regarding the analysis of genetic profiles of humans and other targeted organisms. The application of microarrays with bioinformatic technology is on the threshold of contributing substantially to improved diagnosis and therapeutic management of patients in a clinical setting.

ABBREVIATIONS AND ACRONYMS

bp	base pairs
cDNA	Complementary DNA
G6PD	Glucose-6-phosphate Dehydrogenase
MALDI/TOFMS	Matrix-assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry
mRNA	Messenger RNA
PCR	Polymerase Chain Reaction
RFLP	Restriction Fragment Length Polymorphism
VNTR	Variable Number Tandem Repeat

RELATED ARTICLES

Biomolecules Analysis (Volume 1)

Biomolecules Analysis: Introduction • Fluorescence-based Biosensors

Clinical Chemistry (Volume 2)

Automation in the Clinical Laboratory • Biosensor Design and Fabrication • Capillary Electrophoresis in Clinical Chemistry • DNA Arrays: Preparation and Application • Drugs of Abuse, Analysis of • Molecular Biological Analyses and Molecular Pathology in Clinical Chemistry • Nucleic Acid Analysis in Clinical Chemistry • Phosphorescence, Fluorescence, and Chemiluminescence in Clinical Chemistry • Product Development for the Clinical Laboratory

Forensic Science (Volume 5)

Forensic Science: Introduction • DNA Extraction Methods in Forensic Analysis • Polymerase Chain Reaction in the Forensic Analysis of DNA

Industrial Hygiene (Volume 6)

Metals in Blood and Urine: Biological Monitoring for Worker Exposure

Nucleic Acids Structure and Mapping (Volume 6)

Nucleic Acids Structure and Mapping: Introduction • Capillary Electrophoresis of Nucleic Acids • Comparative Genomics: Differential Display and Subtractive

Hybridization • DNA Molecules, Properties and Detection of Single • DNA Probes • DNA Structures of Biological Relevance, Studies of Unusual Sequences • Fluorescence In Situ Hybridization • Mass Spectrometry of Nucleic Acids • Nuclear Magnetic Resonance and Nucleic Acid Structures • Nucleic Acid Structural Energetics • Optical Mapping in Genomic Analysis • PNA and Its Applications • Polymerase Chain Reaction and Other Amplification Systems • Radiation Hybrid Mapping • Restriction Landmark Genomic and cDNA Scanning • Sequencing and Fingerprinting DNA by Hybridization with Oligonucleotide Probes • Sequencing Strategies and Tactics in DNA and RNA Analysis • X-ray Structures of Nucleic Acids

Electronic Absorption and Luminescence (Volume 12)

Absorption and Luminescence Probes • Fluorescence Imaging Microscopy • Fluorescence Lifetime Measurements, Applications of • Indirect Detection Methods in Capillary Electrophoresis • Surface Measurements using Absorption/Luminescence

General Articles (Volume 15)

Analytical Problem Solving: Selection of Analytical Methods • Multivariate Image Analysis • Quality Assurance in Analytical Chemistry • Quantitative Spectroscopic Calibration

REFERENCES

1. A.E. Garrod, 'Medicine from the Chemical Standpoint', *Lancet*, **ii**, 281–289 (1914).
2. A.E. Garrod, *Inborn Factors in Disease: An Essay*, Oxford University Press, New York, 1931.
3. A.F. Blakeslee, M.R. Salmon, 'Odor and Taste Blindness', *Eugenical News*, **16**, 105–110 (1931).
4. W.W. Weber, *Pharmacogenetics*, Oxford University Press, New York, 1997.
5. A.G. Motulsky, 'Drug Reactions, Enzymes and Biochemical Genetics', *J. Am. Med. Assoc.*, **165**, 835–837 (1957).
6. W. Kalow, *Pharmacogenetics – Heredity and the Responses to Drugs*, W.B. Saunders, Philadelphia, PA, 1962.
7. A.G. Motulsky, 'Pharmacogenetics and Ecogenetics in 1991', *Pharmacogenetics*, **1**, 1–2 (1991).
8. F.P. Guengerich, 'Human Cytochrome P450 Enzymes', in *Cytochrome P450 – Structure, Mechanism, and Biochemistry*, 2nd edition, ed. P.R. Ortiz de Montellano, Plenum Press, New York, 473–517, 1995.
9. D.W. Cox, ' α_1 -Antitrypsin Deficiency', in *The Metabolic and Molecular Bases of Inherited Disease*, eds. C.R. Scriver, A.L. Beaudet, W.S. Sly, D. Valle, 7th edition, McGraw-Hill, New York, 2407–2437, 1995.
10. M.L. Piezza, E.M. Sellers, R.F. Tyndale, 'Nicotine Metabolism Reduces Smoking', *Nature*, **393**, 750 (1998).

11. M. Oscarson, R.A. McLellan, H. Gullsten, Q.-Y. Yue, M.A. Lang, M.L. Bernal, B. Sinues, A. Hirvonen, H. Raunio, O. Pelkonen, M. Ingelman-Sundberg, 'Characterisation and PCR-based Detection of a *CYP2A6* Gene Deletion Found at a High Frequency in a Chinese Population', *FEBS Lett.*, **448**, 105–110 (1999).
12. J. LeWalter, U. Korallus, 'Blood Protein Conjugates and Acetylation of Aromatic Amines: New Findings on Biological Monitoring', *Int. Arch. Occup. Envir. Health*, **56**, 179–196 (1985).
13. R.A. Cartwright, R.W. Glashan, H.J. Rogers, R.A. Ahmad, D.B. Hall, E. Higgins, M.A. Kahn, 'The Role of *N*-Acetyltransferase Phenotypes in Bladder Carcinogenesis: A Pharmacogenetics Epidemiological Approach to Bladder Cancer', *Lancet*, **ii**, 842–846 (1982).
14. E. Beutler, 'Glucose-6-Phosphate Dehydrogenase Deficiency', *N. Engl. J. Med.*, **324**, 169–174 (1991).
15. S. Higuchi, T. Muramatsu, K. Shigemori, M. Saito, H. Kone, M.C. Dufour, T.C. Harford, 'The Relationship between Low K_m Aldehyde Dehydrogenase Phenotype and Drinking Behavior in Japanese', *J. Stud. Alcohol.*, **53**, 170–175 (1992).
16. D.G. McCarver, H.R. Thomasson, S.S. Marteur, R.J. Sokol, T.-K. Li, 'Alcohol dehydrogenase-2,*3 Allele Protects Against Alcohol-related Birth Defects Among African Americans', *J. Pharmacol. Exp. Ther.*, **283**, 1095–1101 (1997).
17. R.K. Ross, J.-M. Yuan, M.C. Yu, G.N. Wogan, G.-S. Qian, J.-T. Tu, J.D. Groopman, Y.-T. Gao, B.E. Henderson, 'Urinary Aflatoxin Biomarkers and Risk of Hepatocellular Carcinoma', *Lancet*, **339**, 943–946 (1992).
18. T.D. Arias, L.F. Jorge, R. Barrantes, 'Uses and Misuses of Definitions of Genetic Polymorphism. A Perspective from Population Pharmacogenetics', *Br. J. Clin. Pharmacol.*, **31**, 117–118 (1991).
19. U.A. Meyer, 'Genotype or Phenotype: The Definition of a Pharmacogenetic Polymorphism', *Pharmacogenetics*, **1**, 66–67 (1991).
20. E.B. Wilson, Jr, *An Introduction to Scientific Research*, McGraw-Hill, New York, Chapter 8, 1954.
21. D. Voet, J.G. Voet, *Biochemistry*, John Wiley, New York, 1990.
22. H.W. Harris, *The Principles of Human Biochemical Genetics*, 3rd edition, Elsevier/North Holland Biomedical Press, New York, 1980.
23. W.W. Weber, 'Influence of Heredity on Human Sensitivity to Environmental Chemicals', *Environ. Mol. Mutagen.*, **25**(Suppl.), 102–114 (1995).
24. J.D. Watson, M. Gilman, J. Witkowski, M. Zoller, *Recombinant DNA*, W.H. Freeman, San Francisco, 1992.
25. K. Muratanai, T. Hada, Y. Yamamoto, T. Kaneko, Y. Shigeto, T. Ohue, J. Furuyama, K. Higashino, 'Inactivation of the Cholinesterase Gene by *Alu* Insertion: Possible Mechanism for Human Gene Transposition', *Proc. Natl. Acad. Sci. USA*, **88**, 11315–11319 (1991).
26. J. Veldscholte, C. Ris-Staplers, G.G.J.M. Kuiper, G. Genster, C. Berrevoets, E. Classen, H.C.J. van Rooij, J. Trapman, A.O. Brinkmann, E. Mulder, 'A Mutation in the Ligand Binding Domain of the Androgen Receptor of Human LNCAP Cells Affects Steroid Binding Characteristics and Response to Antiandrogens', *Biochem. Biophys. Res. Commun.*, **173**, 534–540 (1990).
27. R.C. Skoda, F.J. Gonzalez, A. Demierre, U.A. Meyer, 'Two Mutant Alleles of the Human Cytochrome P-450db1 Gene (*P450C2D1*) Associated with Genetically Deficient Metabolism of Debrisoquine and Other Drugs', *Proc. Natl. Acad. Sci. USA*, **85**, 5240–5243 (1988).
28. N.M. Woolhouse, M. Eichelbaum, N.S. Oates, J.R. Idle, R.L. Smith, 'Dissociation of Co-regulatory Control of Debrisoquine/Phenformin and Sparteine Oxidation in Ghanians', *Clin. Pharmacol. Ther.*, **37**, 512–521 (1985).
29. Q.Y. Yue, L. Bertilsson, M.L. Dahl-Puustinen, J. Sawe, F. Sjoqvist, I. Johansson, M. Ingelman-Sundberg, 'Dissociation between Debrisoquine Hydroxylation Phenotype and Genotype among Chinese', *Lancet*, **ii**, 870 (1989).
30. E. Steiner, L. Iselius, G. Alva, J. Lindsten, F. Sjoqvist, 'A Family Study of Genetic and Environmental Factors Determining Polymorphic Hydroxylation of Debrisoquine', *Clin. Pharmacol. Ther.*, **38**, 394–401 (1985).
31. N. Rothman, G. Talaska, R. B. Hayes, V.K. Bhatnagar, D.A. Bell, V.M. Lakshmi, S.K. Kashyap, M. Dosemeci, R. Kashyap, F.F. Hsu, M. Jaeger, A. Hirvonen, D.J. Parikh, B.B. Davis, T.V. Zenser, 'Acidic Urine pH is Associated with Elevated Levels of Free Urinary Benzidine and *N*-Acetylbenzidine and Urothelial Cell DNA Adducts in Exposed Workers', *Cancer Epidemiol., Biomarkers Prev.*, **6**, 1039–1042 (1997).
32. Q. Zhou, G. Talaska, M. Jaeger, V.K. Bhatnagar, R.B. Hayes, T.V. Zenser, S.K. Kashyap, V.M. Lakshmi, R. Kashyap, M. Dosemeci, F.F. Hsu, D.J. Parikh, B. Davis, N. Rothman, 'Benzidine–DNA Adduct Levels in Human Peripheral White Blood Cells Significantly Correlate with Levels in Exfoliated Urothelial Cells', *Mut. Res.*, **393**, 199–205 (1997).
33. G.H. Goodfellow, K. Durette, D.M. Grant, 'Structure–Function Studies of Human Arylamine Acetyltransferases: The Role of Amino Acids 125 and 127 in Imparting Arylamine Acceptor Substrate Specificity', *First International Workshop on the Arylamine N-Acetyltransferases*, Cairns, Australia, October, 22–24, 1998.
34. N.P. Lang, F.F. Kadlubar, S.L. MacLeod, S. Nowell, J. Massengill, K.E. Anderson, M.D. Gross, 'Influence of NAT-1 Genotype on NAT-2 Phenotype', *First International Workshop on the Arylamine N-Acetyltransferases*, Cairns, Australia, October, 22–24, 1998.
35. H. Wikman, P. Piirinen, K. Kääriä, C. Rosenberg, H. Norppa, A. Hirvonen, 'Role of *NAT1* and *NAT2* Polymorphisms in Individual Susceptibility to Isocyanate-Induced Asthma', *First International Workshop on*

- the Arylamine N-Acetyltransferases', Cairns, Australia, October, 22–24, 1998.
36. H.W. Goedde, D.P. Agarwal, 'Pharmacogenetics of Aldehyde Dehydrogenase (ALDH)', *Pharmacol. Ther.*, **45**, 345–371 (1990).
 37. H. Furuya, P. Fernandez-Salguero, W. Gregory, H. Taber, A. Steward, F.J. Gonzalez, J.F. Idle, 'Genetic Polymorphism of CYP2C9 and its Effect on Warfarin Maintenance Dose Requirement in Patients Undergoing Anticoagulation Therapy', *Pharmacogenetics*, **5**, 389–392 (1995).
 38. W. Kalow, K. Genest, 'A Method for the Detection of Atypical Forms of Serum Cholinesterase. Determination of Dibucaine Numbers', *Can. J. Biochem.*, **35**, 339–346 (1957).
 39. J.R. Playfer, L.C. Eze, M.F. Bullen, D.A.P. Evans, 'Genetic Polymorphism and Interethnic Variability', *J. Med. Genet.*, **13**, 337–342 (1976).
 40. B.N. La Du, S. Adkins, 'Analysis of Serum Para-oxonase/Arylesterase Polymorphism in Some Sudanese Families', in *Ethnic Differences in Reactions to Drugs and Xenobiotics*, eds. W. Kalow, H.W. Goedde, D.P. Agarwal, Alan R. Liss, New York, 87–98, 1986.
 41. D.M. Grant, P. Vohra, Y. Avis, A. Ima, 'Detection of a New Polymorphism of Human Arylamine N-Acetyltransferase NAT1 Using *p*-Aminosalicylic Acid as an In Vivo Probe', *J. Basic Clin. Physiol Pharmacol.*, **3**(Suppl.), 244 (1992).
 42. K.P. Vatsis, W.W. Weber, 'Human N-Acetyltransferases', in *Conjugation–Deconjugation Reactions in Drug Metabolism and Toxicity*, ed. F.C. Kauffman, Springer-Verlag, New York, 1994.
 43. Z. Lu, R. Zhang, R.B. Diasio, 'Dihydropyrimidine Dehydrogenase Activity in Human Peripheral Blood Mononuclear Cells and Liver: Populations Characteristics, Newly Identified Deficient Patients and Clinical Application in 5-Fluorouracil Chemotherapy', *Cancer Res.*, **53**, 5433–5438 (1993).
 44. J. Seidegard, R.W. Pero, 'The Hereditary Transmission of High Glutathione Transferase Activity towards *trans*-Stilbene Oxide in Human Mononuclear Leukocytes', *Human Genet.*, **69**, 66–68 (1985).
 45. J.K. Wiencke, K.T. Kelsey, R.A. Lamela, W.A. Toscano, Jr, 'Human Glutathione-S-transferase Deficiency as a Marker of Susceptibility to Epoxide-induced Damage', *Cancer Res.*, **50**, 1585–1590 (1990).
 46. R. Weinshilboum, 'Methyltransferase Pharmacogenetics', *Pharmacol. Ther.*, **43**, 77–90 (1989).
 47. C.L. Szumlanski, R. Honcel, M.C. Scott, R.M. Weinshilboum, 'Human Liver Thiopurine Methyltransferase Pharmacogenetics: Biochemical Properties, Liver–Erythrocyte Correlation and Presence of Isozymes', *Pharmacogenetics*, **2**, 148–159 (1992).
 48. G. Monaghan, M. Ryan, R. Seddon, R. Hume, B. Burchell, 'Genetic Variation in Bilirubin UDP-gluconosyltransferase Gene Promoter and Gilbert's Syndrome', *Lancet*, **347**, 578–581 (1996).
 49. P.J. Bosma, J.R. Chowdhury, C. Bakker, S. Gantla, A. de Boer, B.A. Oostra, D. Lindhout, G.N.J. Guido, P.M. Jansen, L.M. Peter, O. Elferink, P.J. Ronald, N.R. Chowdhury, 'The Genetic Basis of the Reduced Expression of Bilirubin UDP-glucosyltransferase 1 in Gilbert's Syndrome', *N. Engl. J. Med.*, **333**, 1171–1175 (1995).
 50. M. Patel, B.K. Tang, W. Kalow, '(S)Oxazepam is Inhibited by Ketoprofen and Other Substrates of UGT2B7', *Pharmacogenetics*, **5**, 43–49 (1995).
 51. B.L. Coffman, G.R. Rios, C.D. King, T.R. Tephly, 'Human UGT2B7 Catalyzes Morphine Glucuronidation', *Drug Metab. Dispos.*, **25**, 1–4 (1997).
 52. O. Gotoh, Y. Fujii-Kuriyama, 'Evolution, Structure, and Gene Regulation of Cytochrome P-450', in *Basis and Mechanisms of Regulation of Cytochrome P-450*, eds. K. Ruckpaul, H. Rein, Taylor and Francis, London, 195–243, 1989.
 53. I. Johansson, E. Lundqvist, L. Bertilsson, M.-L. Dahl, F. Sjoqvist, M. Ingelman-Sundberg, 'Inherited Amplification of an Active Gene in the Cytochrome P450 CYP2D Locus as a Cause of Ultrarapid Metabolism of Debrisoquine', *Proc. Natl. Acad. Sci. USA*, **90**, 11825–11829 (1993).
 54. G.K.K. Hershey, M.F. Friedrich, L.A. Esswein, M.L. Thomas, T.A. Chatila, 'The Association of Atopy with a Gain-of-function Mutation in the α -Subunit of the Interleukin-4 Receptor', *N. Engl. J. Med.*, **337**, 1720–1725 (1997).
 55. W.W. Weber, *The Acetylator Genes and Drug Response*, Oxford University Press, New York, 1987.
 56. G.N. Levy, W.W. Weber, 'Interindividual Variability of N-Acetyltransferases in Man', in *Interindividual Variability in Drug Metabolism in Man*, eds. G. Pacifici, O. Pelkonen, Taylor and Francis, London, in press.
 57. D.G. McCarver, R. Byun, R.N. Hines, M. Hicheme, W. Wegenek, 'A Genetic Polymorphism in the Regulatory Sequences on Human CYP2E1: Association with Increased Chloroxazone Hydroxylation in the Presence of Obesity and Ethanol Intake', *Toxicol. Appl. Pharmacol.*, **152**, 276–281 (1998).
 58. E.M. Southern, 'Detection of Specific Sequences Among DNA Fragments Separated by Gel Electrophoresis', *J. Mol. Biol.*, **98**(3), 503–517 (1975).
 59. J.C. Alwine, D.J. Kemp, G.R. Stark, 'Method for Detection of Specific RNAs in Agarose Gels by Transfer to Diazobenzyloxymethyl-paper and Hybridization with DNA Probes', *Proc. Natl. Acad. Sci. USA*, **74**(12), 5350–5354 (1977).
 60. A.R. Wyman, R. White, 'A Highly Polymorphic Locus in Human DNA', *Proc. Natl. Acad. Sci. USA*, **77**(11), 6754–6758 (1980).
 61. A.J. Jeffreys, V. Wilson, S.L. Thein, 'Individual-specific Fingerprints of Human DNA', *Nature*, **316**, 76–79 (1985).

62. A.J. Jeffreys, V. Wilson, S.L. Thein, 'Hypervariable Minisatellite Regions in Human DNA', *Biotechnology*, **24**, 467–472 (1992).
63. Y. Nakamura, M. Leppert, P. O'Connell, R. Wolff, T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin, 'Variable Number of Tandem Repeat (VNTR) Markers for Human Gene Mapping', *Science*, **235**, 1616–1622 (1987).
64. M.S. Ristaldi, M. Pirastu, C. Rosatelli, G. Monni, H. Erlich, R. Saiki, A. Cao, 'Prenatal Diagnosis of Beta-thalassaemia in Mediterranean Populations by Dot Blot Analysis with DNA Amplification and Allele Specific Oligonucleotide Probes', *Prenat. Diagn.*, **9**(9), 629–638 (1989).
65. D.Y. Wu, G. Nozari, M. Schold, B.J. Conner, R.B. Wallace, 'Direct Analysis of Single Nucleotide Variation in Human DNA and RNA Using In Situ Dot Hybridization', *DNA*, **8**(2), 135–142 (1989).
66. E. Kawasaki, R. Saiki, H. Erlich, 'Genetic Analysis Using Polymerase Chain Reaction-amplified DNA and Immobilized Oligonucleotide Probes: Reverse Dot-Blot Typing', *Meth. Enzymol.*, **218**, 369–381 (1993).
67. P. Sutcharitchan, R. Saiki, T.H. Huisman, A. Kutlar, V. McKie, H. Erlich, S.H. Embury, 'Reverse Dot-Blot Detection of the African-American Beta-thalassemia Mutations', *Blood*, **86**(4), 1580–1585 (1995).
68. H. Erlich, T. Bugawan, A.B. Begovich, S. Scharf, R. Griffith, R. Saiki, R. Higuchi, P.S. Walsh, 'HLA-DR, DQ and DP Typing Using PCR Amplification and Immobilized Probes', *Eur. J. Immunogenet.*, **18**(1–2), 33–55 (1991).
69. E.M. Southern, U. Maskos, 'Parallel Synthesis and Analysis of Large Numbers of Related Chemical Compounds: Applications to Oligonucleotides', *J. Biotechnol.*, **35**(2–3), 217–227 (1994).
70. A.C. Pease, D. Solas, E.J. Sullivan, M.T. Cronin, C.P. Holmes, S.P. Fodor, 'Light-generated Oligonucleotide Arrays for Rapid DNA Sequence Analysis', *Proc. Natl. Acad. Sci. USA*, **91**(11), 5022–5026 (1994).
71. A.P. Blanchard, R.J. Kaiser, L.E. Hood, 'High Density Oligonucleotide Arrays', *Biosens. Bioelectron.*, **11**(6/7), 687–690 (1996).
72. M. Schena, D. Shalon, R.W. Davis, P.O. Brown, 'Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray' [see comments]. *Science*, **270**, 467–470 (1995).
73. V.G. Cheung, M. Morley, F. Aguilar, A. Massimi, R. Kucherlapati, G. Childs, 'Making and Reading Microarrays', *Nat. Genet.*, **21**(Suppl.), 15–19 (1999).
74. D.J. Duggan, M. Bittner, Y. Chen, P. Meltzer, J. Trent, 'Expression Profiling Using cDNA Microarrays', *Nat. Genet.*, **21**(Suppl.), 10–14 (1999).
75. D.G. Wang, J.-B. Fan, C.-J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, 'Large-scale Identification, Mapping and Genotyping of Single-nucleotide Polymorphisms in the Human Genome', *Science*, **280**, 1077–1082 (1998).
76. D.J. Lockhart, H. Dong, M.C. Byrne, M.T. Follettie, M.V. Gallo, M.S. Chee, M. Mittmann, C. Wang, M. Kobayashi, H. Horton, 'Expression Monitoring by Hybridization to High-density Oligonucleotide Arrays', *Nat. Biotechnol.*, **14**(13), 1675–1680 (1996).
77. J.G. Hacia, L.C. Brody, M.S. Chee, S.P. Fodor, F.S. Collins, 'Detection of Heterozygous Mutations in *BRCA1* Using High Density Oligonucleotide Arrays and Two-color Fluorescence Analysis', *Nat. Genet.*, **14**(4), 441–447 (1996).
78. J. Hacia, 'Resequencing and Mutational Analysis Using Oligonucleotide Microarrays', *Nat. Genet.*, **21**(Suppl.), 42–47 (1999).
79. M.T. Cronin, R.V. Fucini, S.M. Kim, R.S. Masino, R.M. Wespi, C.G. Miyada, 'Cystic Fibrosis Mutation Detection by Hybridization to Light-generated DNA Probe Arrays', *Hum. Mutat.*, **7**(3), 244–255 (1996).
80. V. Chan, D.J. Graves, S.E. McKenzie, 'The Biophysics of DNA Hybridization with Immobilized Oligonucleotide Probes', *Biophys J.*, **69**(6), 2243–2255 (1995).
81. L. Wodicka, H. Dong, M. Mittmann, M.H. Ho, D.J. Lockhart, 'Genome-wide Expression Monitoring in *Saccharomyces cerevisiae*', *Nat. Biotechnol.*, **15**(13), 1359–1367 (1997).
82. D. Shalon, S.J. Smith, P.O. Brown, 'A DNA Microarray System for Analyzing Complex DNA Samples Using Two-color Fluorescent Probe Hybridization', *Genome Res.*, **6**(7), 639–645 (1996).
83. J.G. Hacia, S.A. Woski, J. Fidanza, K. Edgemon, N. Hunt, G. McGall, S.P. Fodor, F.S. Collins, 'Enhanced High Density Oligonucleotide Array-based Sequence Analysis Using Modified Nucleoside Triphosphates', *Nucl. Acids Res.*, **26**(21), 4975–4982 (1998).
84. U. Maskos, E.M. Southern, 'Oligonucleotide Hybridizations on Glass Supports: A Novel Linker for Oligonucleotide Synthesis and Hybridization Properties of Oligonucleotides Synthesised In Situ', *Nucl. Acids Res.*, **20**(7), 1679–1684 (1992).
85. W.I. Wood, J. Gitschier, L.A. Lasky, R.M. Lawn, 'Base Composition-independent Hybridization in Tetramethylammonium Chloride: A Method for Oligonucleotide Screening of Highly Complex Gene Libraries', *Proc. Natl. Acad. Sci. USA*, **82**(6), 1585–1588 (1985).
86. T.T. Nikiforov, R.B. Rendle, M.L. Kotewicz, Y.H. Rogers, 'The Use of Phosphorothioate Primers and Exonuclease Hydrolysis for the Preparation of Single-stranded PCR Products and Their Detection by Solid-phase Hybridization', *PCR Meth. Appl.*, **3**(5), 285–291 (1994).
87. T.T. Nikiforov, R.B. Rendle, P. Goelet, Y.H. Rogers, M.L. Kotewicz, S. Anderson, G.L. Trainor, M.R. Knapp, 'Genetic Bit Analysis: A Solid Phase Method for Typing Single Nucleotide Polymorphisms', *Nucl. Acids Res.*, **22**(20), 4167–4175 (1994).

88. M. Chee, R. Yang, E. Hubbell, A. Berno, X.C. Huang, D. Stern, J. Winkler, D.J. Lockhart, M.S. Morris, S.P. Fodor, 'Accessing Genetic Information with High-density DNA Arrays', *Science*, **274**, 610–614 (1996).
89. C.F. Edman, D.E. Raymond, D.J. Wu, E. Tu, R.G. Sosnowski, W.F. Butler, M. Nerenberg, M.J. Heller, 'Electric Field Directed Nucleic Acid Hybridization on Microchips', *Nucl. Acids Res.*, **25**(24), 4907–4914 (1997).
90. R.G. Sosnowski, E. Tu, W.F. Butler, J.P. O'Connell, M.J. Heller, 'Rapid Determination of Single Base Mismatch Mutations in DNA Hybrids by Direct Electric Field Control', *Proc. Natl. Acad. Sci. USA*, **94**(4), 1119–1123 (1997).
91. J. Cheng, E.L. Sheldon, L. Wu, A. Uribe, L.O. Gerrue, J. Carrino, M.J. Heller, J.P. O'Connell, 'Preparation and Hybridization Analysis of DNA/RNA from *E. coli* on Microfabricated Bioelectronic Chips', *Nat. Biotechnol.*, **16**(6), 541–546 (1998).
92. H. Koster, K. Tang, D.J. Fu, A. Braun, D. van den Boom, C.L. Smith, R.J. Cotter, C.R. Cantor, 'A Strategy for Rapid and Efficient DNA Sequencing by Mass Spectrometry', *Nat. Biotechnol.*, **14**(9), 1123–1128 (1996).
93. D.P. Little, A. Braun, M.J. O'Donnell, H. Koster, 'Mass Spectrometry from Miniaturized Arrays for Full Comparative DNA Analysis', *Nat. Med.*, **3**(12), 1413–1416 (1997).
94. J.A. Ferguson, T.C. Boles, C.P. Adams, D.R. Walt, 'A Fiber-optic DNA Biosensor Microarray for the Analysis of Gene Expression', *Nat. Biotechnol.*, **14**(13), 1681–1684 (1996).
95. K.L. Michael, L.C. Taylor, S.L. Schultz, D.R. Walt, 'Randomly Ordered Addressable High-density Optical Sensor Arrays', *Anal. Chem.*, **70**(7), 1242–1248 (1998).
96. T.M. Cox, 'Aldolase B and Fructose Intolerance', *FASEB J.*, **8**, 62–71 (1994).
97. T.M. Cox, 'Iatrogenic Deaths in Hereditary Fructose Intolerance', *Arch. Dis. Child.*, **69**, 413–415 (1993).
98. M. Ali, T.M. Cox, 'Diverse Mutations in the Aldolase B Gene that Underlie the Prevalence of Hereditary Fructose Intolerance', *Am. J. Hum. Genet.*, **56**, 1002–1005 (1995).
99. R. Ayeshe, R.L. Smith, 'Genetic Polymorphism of Trimethylamine N-Oxidation', *Pharmacol. Ther.*, **45**, 387–401 (1990).
100. C.T. Dolphin, A. Janmohamed, R.L. Smith, E.A. Shephard, I.R. Phillips, 'Missense Mutation in Flavin-containing Mono-oxygenase 3 Gene, *FMO3*, Underlies Fish-odour Syndrome', *Nat. Genet.*, **17**, 491–494 (1997).
101. W. Burke, E. Thomson, M.J. Khoury, S.M. McDonnell, N. Press, P.C. Adams, J.C. Barton, E. Beutler, G. Brittenham, A. Buchanan, E.W. Clayton, M.E. Cogswell, E.M. Meslin, A.G. Motulsky, L.W. Powell, E. Sigal, B.S. Wilford, F.S. Collins, 'Hereditary Hemochromatosis: Gene Discovery and its Implications for Population-based Screening', *J. Am. Med. Assoc.*, **280**, 172–178 (1998).
102. F.S. Collins, 'Cystic Fibrosis: Molecular Biology and Therapeutic Implications', *Science*, **256**, 774–779 (1992).
103. W.W. Weber, 'Populations and Genetic Polymorphisms', *Mol. Diagn.*, **4**(4), 299–307 (1999).
104. D. Marez, M. Legrand, N. Sabbagh, J.-M. Lo Guidice, J.-M. Spire, J.-J. Lafitte, U. Meyer, F. Broly, 'Polymorphism of the Cytochrome P450 CYP2D6 Gene in an European Population: Characterization of 48 Mutations and 53 Alleles, their Frequencies and Evolution', *Pharmacogenetics*, **7**, 193–202 (1996).
105. C. Sachse, J. Brockmüller, S. Bauer, I. Roots, 'Cytochrome P450 2D6 Variants in a Caucasian Population: Allele Frequencies and Phenotypic Consequences', *Am. J. Hum. Genet.*, **60**, 284–295 (1997).
106. M. Ingelman-Sundberg, 'The Gerhard Zbinden Memorial Lecture. Genetic Polymorphism of Drug Metabolizing Enzymes. Implications for Toxicity of Drugs and Other Xenobiotics', *Arch. Toxicol. Suppl.*, **19**, 3–13 (1997).
107. G.R. Wilkinson, F.P. Guengerich, R.A. Branch, 'Genetic Polymorphism of *S*-Mephenytoin Hydroxylation', *Pharmacol. Ther.*, **43**, 53–76 (1989).
108. G. Ibeanu, J. Blaisdell, B.I. Ghanayem, C. Beyeler, S. Benhamou, C. Bouchardy, G.R. Wilkinson, P. Dayer, A.K. Daly, J.A. Goldstein, 'An Additional Defective Allele, *CYP2C19**5, Contributes to the *S*-Mephenytoin Poor Metabolizer Phenotype in Caucasians', *Pharmacogenetics*, **8**, 129–135 (1998).
109. F.S. Collins, 'Shattuck Lecture – Medical and Societal Consequences of the Human Genome Project', *N. Engl. J. Med.*, **341**, 28–37 (1999).
110. S. Sindrup, K. Brosen, 'The Pharmacogenetics of Codeine Hypoalgesia', *Pharmacogenetics*, **5**, 335–346 (1995).
111. R.F. Tyndale, K.P. Droll, E.M. Sellers, 'Genetically Deficient CYP2D6 Metabolism Provides Protection Against Oral Opiate Dependence', *Pharmacogenetics*, **7**, 375–379 (1997).
112. D. Wu, V. Otton, T. Inaba, W. Kalow, E. Sellers, 'Interactions of Amphetamine Analogs with Human Liver CYP2D6', *Biochem. Pharmacol.*, **53**, 1605–1612 (1997).
113. P. Dalen, C. Fregnell, M.-L. Dahl, F. Sjöqvist, 'Quick Onset of Severe Abdominal Pain after Codeine in an Ultrarapid Metabolizer of Debrisoquine', *Ther. Drug Monit.*, **19**, 543–544 (1997).